OBITUARY NOTICE

DAVID KATZ

David Katz, Emeritus Professor of Psychology at Stockholm University, died on 2 February at Stockholm. In the last two years of his life, having triumphed over many difficulties, he had the satisfaction of presiding as host over the International Congress of Psychology and of lecturing at the University of Hamburg, after having left Rostock University in 1933 as a result of Nazi persecution.

Since with his usual care he has recorded in his contribution to *A History of Psychology in Autobiography* (vol. IV)[1] the titles and dates of his publications, with explanatory comments, the present appreciation will include only works illustrating his salient characteristics.

To him, it can truly be said, nothing human was alien. To him, in a high degree, had been given the rare ability to observe objects and situations as freshly as if he had never heard of a description or explanation of them. This may explain why, though he attracted many research students, he founded no school of thought and, I believe, never joined one. Even his interest in the Gestalt movement, though he was a close friend of its leaders, was peripheral and critical, as his book *Gestalt Psychology*[2] shows. He maintained that sensation-qualities, the symbol-concept, isomorphism and differential psychology all present problems relatively unrelated to those of configuration. Though his knowledge of the older sciences was considerable, he believed that a psychologist ought never to feel constrained to present his findings in their terminology, perhaps suspecting that it would be salutary for practitioners of the older sciences to recognize this. He always tried to understand facts directly given, without presupposition or prejudice.

Of this sophisticated-naïve approach I recall many examples in his casual questions to me. 'How do you, a visualizer, set about looking for a lost article? If you succeed or fail, is it due in any way to your visualization?' 'Why haven't the dog-loving English supported more research into canine hysteria? When this dog [a victim] was jumping in the air and snapping at nothing, could he have had visual hallucinations?' 'Why do most text-books of psychology in English, French, German, Indian and perhaps even Japanese, contain almost identical chapter-headings?' (this was truer in 1933 than now). Concerning social facts, his observations (he loved this philosophical phrase) were individualistic. English university students do less work than their German opposite-numbers, but they are not so blind to world-events. He held—perhaps this typified an occasional streak of old-fashionedness—that this country's behaviour after King Edward VIII's abdication demonstrated its political maturity.

His psychological interests were never constricted by convention or fashion. Though composite photographs had been used by Galton before he was born, he was employing them, and suggested their further use (in a letter to me) up to a few months before his death.

Of the relation of psychology to medical problems, except prosthesis and massage, he seldom spoke. Perhaps deliberately he did not commit himself in public about the pros

[1] Oxford University Press.

[2] 1951. London: Methuen.

and cons of the issues raised by psycho-analysis. I believe he had not been analysed and sometimes wonder whether such treatment might have decreased his delight in perceiving and thinking unusually, and extinguished the pixie-like twinkle with which he used to share such experiences with sympathetic hearers. I accompanied him to a police station to register his arrival in Manchester. Finding himself at a counter labelled 'Aliens, Explosives, Entertainments', he delighted the lady official with a smile and 'Good-morning; I am all three of these!' Interpreting this remark in its context, Freud might have unearthed a rich condensation of mechanisms.

An excellent experimentalist, Katz quoted with approval an astronomer's remark about a new giant telescope; 'the most important part of the apparatus is the experimenter sitting in front of the lens'. His comments 'Psychology, as carried out in Göttingen as well as in other universities, needed psychologising'; 'The danger of misinterpretation of the relation between psychology and statistics existed in the older experimental psychology and appears again to-day, although on a higher level. . . .Psychological insight cannot be replaced by statistics, however specialized', show a resistance to seduction by method for method's sake.

His Presidential Address at Stockholm in 1951, on 'The Margin of Safety' showed no diminution of his freshness of outlook. Plenty of speculation was supported by plenty of evidence. The reciprocal innervation of antagonistic muscles, anticipation of accidents, insurance-policies, attitudes of optimism and pessimism, even perhaps a hint on care in washing-up; all these heterogeneous themes fell into place under his synoptic glance.

His book *Modes of Appearance of Colours and their Modification through Individual Experience* (the English title *The World of Colour* is an over-simplification) asked 'Why not study the psychological phenomena of colour directly, instead of restricting oneself to problems which can be handled just as well or even better by a sensory physiologist?' One should begin, he argued, not with remote and isolated colour-experiences (many people die without ever having looked through a spectroscope) but actual appearances, as film-, surface- and volume-colours; and the facts of colour constancy should be recognized.

He was always interested in fellow-creatures, including dogs, that through ill-luck had lost a sense or a limb. In 1918 he worked on the psychology of prosthesis, fitting amputated parts with the best artificial substitutes. This raises many psychological questions; the subject was still interesting technicians and physicians at Berkeley when he was there in 1950. With Révész he published a study of musical enjoyment by a totally deaf man, showing that the vibratory sense, a bridge between the pressure-sense and the auditory sense, can function vicariously for the absent auditory sense.

Conversations with Children (1927) broke new ground by choosing, as material for analysis, not words or sentences but conversation, often arising out of his two little boys' spontaneous questions. In this work he was greatly helped, as so often in his life, by his wife Rosa. One inquiry, unrecorded I believe in the book, he would recount with glee, 'Daddy, were you the David that killed Goliath?'

In a few lines he has told the story of persecution as it affected him and his family. The Government of Mecklenburg declared the chair that he occupied to be unnecessary, and suspended the appointment, in spite of the fact that psychology had become the largest department of the University. In the autumn of 1933 he accepted an invitation to come to the University of Manchester as Honorary Research Fellow. During his

1½ years' stay, he brought a never-failing supply of new ideas, and made many friends, not only in the psychological laboratory. He carried on research which led to the publication of his paper on 'The Tongue as a Primitive Sense-Organ', and lectured on the psychology of the blind and of animals. He left Manchester to work in University College, London, at Sir Cyril Burt's invitation and at the Royal Zoological Gardens, where Dr Julian S. Huxley gave him and his wife facilities to work on the vision and feeding behaviour of monkeys. His book *Animals and Men* will soon be published in a revised edition in the Pelican Series. Characteristically, his interests never remained inside the conventional triangle, rat-monkey-man, and recent investigators have followed his broad-minded example.

Katz observed people; he liked most of them and loved many of them. Can one study music without loving it in some way? Perhaps this is true of the proper study of Mankind.

T. H. PEAR.

HOMEOSTASIS, NEEDS AND VALUES[1]

By C. A. MACE

Department of Psychology, Birkbeck College, University of London

I. *Introduction* (pp. 200–202). II. *Two models of causality in nature* (pp. 202–204). III. *Homeostasis and related concepts* (pp. 204–206). IV. *The philosophical theories of value and the 'semantic' problem* (pp. 206–208). V. *Biological concepts of value* (pp. 208–209). VI. *Value versus fact* (pp. 209–210).

I. Introduction

'Science', it is said, 'is not concerned with values. When applied to human affairs it can assist in devising means: it has nothing to do with ends.' These assertions are apt to pass unchallenged, but starkly confronting them are the following facts. Medical science aims at promoting 'health' and preventing 'disease'. Biologists speak of 'survival value', of biological needs, even of biological 'norms'. The social sciences are concerned with the welfare of man in society. But it is in psychology that one meets the sharpest contradiction between practice and profession.

The educational psychologist behaves as if he knows what is good for the child. The psychotherapist behaves as if he knows what is good for the patient. The occupational psychologist conducts himself as though he knows what is good for the workers. *And so in fact they do.* When they say, as they often do when challenged on the point, that their science is not concerned with values, they are, it would seem, merely being inconsistent. This inconsistency is not peculiar to applied psychology. Many of the most general of the theoretical concepts of psychology are evaluative. 'Adjusted' is covertly evaluative. So is 'integration'. So, too, is 'co-ordination'. So, even, is the concept of a 'learned response'. All these concepts require to be defined in terms of adaptation to an end, or of fitness for a purpose. Each of these concepts involves a reference to a 'goal' or purpose; and in most cases the purpose is one which the psychologist implicitly endorses. And the responsibility for clearing up the present confusion in this situation rests upon the theoretical psychologist. It is for him to construct the theory which would make the psychologist's professions consistent with his practice. Nor should it be difficult to do so. The inconsistency has arisen not from within psychology itself but from an uncritical acceptance by psychologists of one particular reading of the philosopher's distinction between the 'positive' and 'normative' sciences. On this reading of this distinction 'values' are opposed to 'facts'. Within psychology itself another reading is possible, that in which the important distinction is between *what is in fact desired* in the ordinary way and *what would in fact be desired* under conditions more favourable to considered choice. This interpretation need not be restricted to any narrow school. It is all too easy to exaggerate the conflicts between the schools of psychology. But when due allowance is made for the confusion of tongues in the psychologists' tower of Babel there is found to be an impressive measure of agreement concerning the nature of man and the goals of his behaviour.

[1] An abstract from the author's Presidential Address to the British Psychological Society delivered at Nottingham on 11 April 1953, under the title 'A psychologist's approach to the theory of values'.

Man can be usefully described as a collocation of drives, propensities or needs, more or less specific and more or less variable in their modes of satisfaction. Initially, these drives are lacking in coherence, but they tend to fall into a sort of hierarchical system, and by being ordered in this way to attain a measure of consistency. A conception of this sort was first presented in a fairly detailed way in Bishop Butler's *Sermons on Human Nature*. Butler describes human nature as a three-tiered hierarchy. At the base of the pyramid lie the particular impulses; at the second level the two regulative principles of Cool Self-love and Benevolence; and at the apex the supreme controlling agent Conscience. The schema is reproduced with remarkable fidelity in the system of McDougall with his three-tiered hierarchy of instincts, sentiments and the master sentiment of Self Regard. Nor is the pattern very different in Freud, who has his own three-tiered hierarchy of Id, Ego and Super-Ego. If McDougall may be described as Butler in modern dress, Freud might be described as Butler in fancy dress. But the analogies though instructive must not be pressed too far.

All accounts of this general type would seem to assert or imply that 'the good for man' consists in the mutual adjustment of his needs and in their harmonious satisfaction, in the attainment, that is, of the greatest satisfaction on the whole and in the long run. In principle, the theory is clear and simple, but it requires to be articulated in greater detail. There would appear to be three ingredients in the doctrine: (i) a definition of the primary units in the hierarchical system, (ii) an account of the manner in which these units may be modified and perhaps of the way in which new units of the required type may emerge, (iii) an account of their 'integration'.

It does not matter much what we call the basic units—'primary instincts', 'propensities', 'drives', 'needs', 'ergs', 'action potentials' or any of the other names that have been suggested. So far as they may in fact differ they can be regarded as functionally equivalent concepts. All enter in the same way into the theoretical system in which they are employed.

There is some convenience in starting from the concept of a 'need'. It is not a technical term, it is in common use, and it is perhaps less ambiguous than any of the alternatives that might be suggested. It might almost be claimed that the idea of a 'need' is both familiar and luminously clear. We all have many needs, we all experience needs, is there anything odd in that? In point of fact there is something odd in the idea, especially for the natural scientist working within the confines of traditional concepts. The idea carries an aroma of brimstone or what in this connexion is much the same thing, it is suggestive of 'teleology' and evaluation. This indeed is so. 'Need' *is* a teleological concept, and to establish its position in a reputable biology has reopened a number of questions which many had supposed to be closed. This is, of course, a very long story, going back to Aristotle, and his idea of 'final causes'; but a more convenient point of time at which the story may be taken up is that at which the philosopher Spinoza presented to the world his concept of *Conatus*. 'Everything that is in itself', said Spinoza, 'endeavours to persist in its own being.' The suggestion seems to be that everything *whatever* resists change, and if forcibly changed tends to revert to its prior state. Spinoza was mistaken in supposing this to be true of everything whatever, but it would seem that it *is* true of living things. As Broad has observed,'modern physiology agrees entirely with this doctrine of Spinoza's and its researches have established it in much greater detail than Spinoza could have dreamt of'. Evidence in support of it can in fact be adduced from all the biological

sciences. The increasing use by biologists of a principle equivalent to that of Spinoza's *Conatus* means that in effect scientists have now two contrasted models for the description of causality in Nature.

II. Two models of causality in nature

The first is the model of the inert body, of a thing that does not move until it is pushed, of a thing that does not change in any way until something else acts upon it. The paradigm of this is the classical billiard ball.

Contrasted with this is the model of the inherently active system—the sort of thing that behaves and continues to behave in a characteristic way until something interferes to change its behaviour. The distinction might be illustrated by reference to the contrast between material things as they appear in ordinary perception and material things as described by physics. The billiard ball sensibly appears to be, in itself, an unchanging, homogeneous inert body, whereas physics describes it as a system of atomic or subatomic elements moving or changing in characteristic ways, and continuing so to move or change until something comes along to change the ways in which they move and change.[1] Both accounts may well be correct since properly stated they are not inconsistent; but it also well may be that the thing as it appears is a perceptual artefact—a perceptual artefact in the sense in which the phi-phenomenon is so. It well may be the case that the second model is *par excellence* the model of all scientific descriptions, and that the first is *par excellence* the model of prescientific thought. Be this as it may, many examples of the use of this second kind of model could be cited. Of greater interest to psychologists is one that is found in contemporary neurology. The traditional conception of the nervous system portrayed it as a structure that remained inert in the absence of stimulation, like a dead electrical circuit before the current is switched on. To this is opposed the idea of the nervous system as an *active* system which continues to tick over in a characteristic way in the absence of stimulation. On this view the function of the stimulus is not to induce activity in a structure previously inert but to modify an activity already in progress.[2]

On the nature of a 'need'

Now, this second type of model would seem to be the more appropriate for the understanding of needs. A system acting under the stress of a need can be described as one which has the following characteristics:

It is a system for which there is a certain possible state such that when the system is in that state it is quiescent (or free from stress or tension), when this state is absent or disturbed the system becomes active and, in a variety of ways, accidents apart, it continues to be active until that state is established or restored.

[1] Again it does not matter, in the present context, what account is given of the nature of the 'elements' in the system. Different systems of physics will have functionally equivalent concepts.

[2] The thesis with which Hebb opens his argument in the *Organization of Behaviour* turns on a distinction of this kind. It is, however, of historical interest to note that a similar function is assigned to peripheral stimulation in Descartes' account of reflex action. The peripheral stimulus merely helps to influence the course of the animal spirits on their way from the heart to the executive organs.

I am indebted to Miss Thelma Veness for drawing my attention to W. A. Murray's distinction between 'proaction' and 'reaction' ('Towards a Classification of Interaction' in *Towards a General Theory of Action*, ed. T. Parson, and E. Shils, Harvard University Press, 1951) which appears to be based upon the same fundamental dichotomy.

We may say of such a system that its activity is 'goal directed', and that the state in which it would be quiescent or free from tension is its end or goal. This is a definition of the expression 'goal directed'. Alternatively, we may say that the system has a *need*, and that what it needs is defined as the 'goal' of its activity.

This formula defines the abstract and limiting case of a system with a single need. We can conceive a more complex system for which there are n such states; and if that system were a living thing we could say that those n states were the ends or goals of its life. If the system in question could introspect and by introspection could define its goals prior to their attainment then we could say of such a system that it 'knows what it wants'. This, too, is *only* a definition, but it is a definition which may have some importance, since it implies conditions under which an introspective report might be validated. If, for example, the System says 'I want x' and then giving it x does not produce the appropriate kind of quiescence, this is evidence that the system does not really know what it wants. The system does not *really* want x. It merely 'thinks it does'.

This account of a need is much too abstract. It can, however, be put much more concretely. Take first a simple, hypothetical case of a system under the stress of a need. Suppose that you had a billiard ball that behaved like a rabbit. Suppose, that is to say, that when you placed this billiard ball in the middle of a perfectly level billiard table it rolled into one of the corner pockets. When you replaced it in the middle of the table and tilted the table so as to raise the level of the pocket, the ball rolled *uphill* into the pocket. When you placed an obstacle in its way it made a detour round the obstacle and again into the pocket. You then put it in the centre of a complicated maze, and after a certain amount of apparently random rolling in and out of blind alleys it escaped from the maze and away again into the pocket. And so for all other variations. The ball is inert or quiescent only when it is in its pocket. Under all other conditions it is 'spontaneously' active, variously active, and it continues to be spontaneously active except when it occupies its pocket. What in these circumstances would we be inclined to say? We should say, shouldn't we, that the ball behaves as though for all the world it were alive, and that it had 'an instinct' or drive impelling it to its pocket. In the terminology here adopted it behaves as if it acted under the stress of a need to occupy a pocket.

Of course, billiard balls do not behave like that. But the importance of this ridiculous hypothetical case is that it draws attention to the fact that we can properly say that a system is active under the stress of a need without introspective evidence, and without any evidence regarding the internal constitution of the system in question. We say this solely on the evidence of its observed molar or macroscopic behaviour.[1]

A billiard ball does not behave like a rabbit, but the rabbit does. So consider the rabbit. The point of special interest in the rabbit or other organism is the great variety of its activities which are goal directed in the sense defined and the great variety in the ways in which any particular goal may be attained. An impressive survey of the varieties of goal-directed behaviour has been given in E. S. Russell's *Directiveness of Organic Activities*. This survey is impressive not only for its range but also for the generality with which the principle is displayed. Three types of goal-directedness are distinguished by Russell—

[1] Again, as a passing historical note, it was McDougall, the arch anti-behaviourist, who was the first to give (without perhaps fully realizing what he was doing) a 'behaviourist' analysis of purposive behaviour. The marks of behaviour set out in his *Outline of Psychology* are broadly those contained in the above description of a system under the stress of a need.

the behavioural, the physiological and the morpho-plastic. The need for warmth, for example, can be satisfied by increased vigour of bodily activity, by the physiological mechanism for the regulation of body temperature or, as a longer term policy, by the growth of hair. If we extend the survey further to cover the relevant facts of psychology and the social sciences, other forms of behaviour directed to the maintenance of an optimum temperature can be distinguished—such as the making of clothes and the construction of thermally insulated buildings.

III. Homeostasis and related concepts

The most familiar form of physiological goal-directed activity is that which has come to be described as 'homeostasis'. Of outstanding importance in the development of this concept is that in it we begin to pass from description only to explanation, and the sort of explanation that can be given in terms of—to use Sherrington's expression—'understandable chemistry'. In fact the explanation of the behaviour of the simplest form of homeostat can be given in terms of understandable physics. It is not surprising, then, that homeostasis has tended to become the model for the interpretation of every kind of goal-directedness.

But homeostasis is not enough. This concept is of its greatest value when it is restricted to the case in which (1) the norm is defined in terms of some internal condition of the organism, (2) this norm is merely maintained or restored and (3) it relates to some specific need as contrasted with the general welfare of the body or the personality as a whole.

In this statement the term 'norm' is used only in the sense in which it is strictly synonymous with that of the term 'end' or 'goal' as above defined. But at this point it is convenient to mention a point on which misunderstanding might arise. A norm or end state need not be conceived as a *static* or *unchanging* condition of the organism. On the contrary, we may here again prefer the model of the *inherently active system* to the model of the *inherently inert system.* When a homeostat restores the normal state that state *may itself be a state of change.* The cycle is: the system ticks over in a normal way, it is disturbed, and it changes its activity until it has restored the condition of ticking over in its normal way. And even to describe the normal state as one of ticking over may be to write down the complexity of the normal activity presupposed. Certainly as we pass from the simpler so-called 'biological' needs to the more complex 'psychological' and social needs the more frequently shall we describe the so-called end *states* in terms of optimal rates and forms of change. There is a very old dictum to the effect that 'All consciousness is consciousness of change'. It might be paralleled by the dictum 'All life is living through change.' Neither psychology nor biology nor sociology can find much use for the concept of an inert substance or system in a static state. With this general understanding we can continue to use the expressions 'end state' and 'norm'. And using these expressions with this understanding we can proceed to extend or generalize the concept of homeostasis in three ways.

The first extension would cover the case in which what is maintained or restored is not so much an internal state of the organism as some relation of the organism to its environment. This would take care of the facts of adaptation and adjustment, including adjustment to the social environment. This would go part of the way towards a clearer and more precise analysis of the vital concept of a 'well adjusted personality'.

The second extension would cover the case in which the goal, end or norm is some state or relation which has never previously been experienced. There is clearly no reason to suppose that every process of the homeostatic type consists in the maintenance or restoration of a norm. There is no reason whatever to suppose that the process always begins in a state of equilibrium which is then disturbed. Life may *start* in disequilibrium and in maladjustment, so that when the goal-directed process reaches its destination a state or relation is established that has never obtained before. There are, at any rate, many cases in which we require the concept of homeostasis to be extended so that it may apply not only to the *restoration* of an equilibrium but also to the discovery of new equilibria.

In social processes the two forms of homeostasis correspond to the two main forms of Utopianism—the nostalgic and the prospective—the one longing for the restoration of a golden age of the past, the other living in hope of a New Jerusalem.

Growth or maturation is an interesting case. Maturation would seem to be a goal-directed process in a sense in which senescence is not. The child has a *need* to grow up, a need in the quite specific sense defined. Peter Pan, the mythological exception, was a deviant from the norm. The organism is an activity system displaying a characteristic cycle of activity which comes to its characteristic termination in maturity. There is thus a homeostatic theory, in the extended sense, of the process of maturation. It might be supposed that maturation is genetically the restoration or reproduction of a prior state—the state of the parent. But even this is not quite correct. Later generations do not merely reproduce the characteristics of earlier generations. We are most like our parents not at maturity but in old age. This is not a fact of maturation, but falls under the very different concept of 'regression'.

Then there are the facts of learning. Three sorts of cases may be noted under this general rubric in which, in the end state of a goal-directed process, something is discovered that has never been experienced before or in which new goals are created. The first is the simple case of cognitive discovery. Observational search by means of the senses is a goal-directed process. So, too, is the process of thought. The facts are set out in the textbooks in the chapter on attention, the chapter which deals with curiosity and the chapter entitled 'The Higher Processes of Thought'. The central facts are that tensions are set up when we are presented with something incomplete, unclear or unfamiliar. A task is prescribed, a problem presented or a question posed. Processes ensue which terminate in the completion of the task, the solving of the problem or the answering of the question. Something is discovered and in discovery tension is relieved. A generalized theory of homeostasis would then cover the perceptual tensions, and by the same token there is awaiting formulation in detail a homeostatic theory of thought.[1]

It is clear that in the course of evolution, as in the course of individual development, new goal-directed systems come into being. Man is continually finding new things to go for. Goals emerge through sensory and through motor differentiation. The development of vision and hearing has created the possibility of the major arts. Every sense has in principle its art. For every modality of sense there is a modality of need. So, too, with the development of skills. Movements and activities can be satisfying in themselves and

[1] Plato's theory that what we seem to 'discover' is in fact reminiscence or rediscovery is of interest in that, if it were true, discovery could be described as a homeostatic process of the simple type. But, of course, it would be agreed that on this point he was quite fantastically mistaken.

irrespective of their more remote effects. There is clearly a need for movement as such and many needs for specific types of bodily activity. These are paralleled by the needs for what we call 'activity of mind'.

New goals emerge through the well-known mechanisms of 'association', 'conditioning' and the 'transference of affect'. These, too, may be regarded as functionally equivalent concepts. They do corresponding jobs in the several types of learning theory. For the study of values the most important branch of learning theory is that which accounts for acquired tastes. The child dislikes, say, spinach. Skilfully handled, he agrees to eat it 'because it is good for him'. The miracle happens. After a time he likes spinach. Other miracles may occur. He may in the same way acquire a taste for Latin or a need for commendation. These are acquired needs. He becomes, so to speak, homeostatic in respect of his intake of spinach, and homeostatic in respect of his output of good behaviour.

The principle would seem to be, formally stated: If a system is goal-directed (in the sense defined) to a state E and this state is attained (in greater or less degree) in the setting of, or with the concomitance of, a state C, then the system will become goal-directed to C, and under certain conditions may cease to be goal-directed to E.

The third way in which the concept of homeostasis can be extended is that in which it would cover second and higher order needs. The various original and acquired needs tend to become mutually adjusted. To this, education and therapy make their distinctive contributions, but it would appear that the mutual adjustment of goal-directed tendencies is in some degree spontaneous. It is found at levels of life below those at which formal education and therapy can operate. An organism, said Aristotle, is a system of organs mutually adjusted to the ends of the organism as a whole. The modern biologist might say that an organism is a system of directed activities mutually adjusted to the goals of the organism as a whole—survival, growth, reproduction, etc. And the psychologist might make a similar statement concerning the nature of a personality. Such a conception would seem to be implied by those, like Hadfield, who define mental health as a state in which full and free expression is given to all native and acquired potentialities in harmony with one another by being directed towards a common end or aim of the personality as a whole. More briefly it is the state in which the greatest satisfaction would be given to the *total system* of original and acquired needs.

This, then, in the sketchiest of outline, is one psychological approach to the theory of values. It rests upon: (i) a generalized concept of a goal-directed tendency provisionally described as a 'need'; (ii) the recognition of the ever expanding nature of the system of needs and the orientation of learning theory under the paradigm of the acquired taste; (iii) a concept of 'the integration of personality' defined in terms of the mutual adjustment of needs.

IV. The philosophical theories of value and the 'semantic' problem

This is not the place in which to discuss the relations of the approach here outlined to the traditional philosophical theories, but neither can these relations be entirely ignored. One of the larger issues as philosophers state it is sufficiently familiar:

When we make a value judgement, for example that 'This is good', it has seemed to some that there is being attributed to whatever is judged to have that value a certain distinctive property that it would possess regardless of the emotions, the desires, the

needs of those who pass this judgement or the emotions, desires or needs of anybody else. On the other hand, it has seemed to others that the judgement states or expresses a fact about the emotions, desires or needs either of the person making the judgement or the emotions, desires or needs of the members of some 'reference group'. On the first view, value is 'absolute' and 'objective'; on the second, it is relative and subjective. The issue is in part, of course, 'a question of words', a 'semantic' question or a question of 'conceptual analysis'; but this is the sort of 'philosophical' question in which the scientist, whether he likes it or not, has to take some interest. If he talks at all, he must be prepared to say what he means when he uses a word the meaning of which may be ambiguous. 'Good' is a word with many meanings, some are trivial, others important. In some of these meanings whether a thing is or is not good is not a question of empirical fact. In other senses it is. There are senses of 'good' in which whether a thing is or is not good can only be decided by reference to the observable facts.

We may note a certain difference between the philosopher's reaction, and the scientist's reaction to ambiguity and plurality of meanings. There are philosophers of many different kinds. In the present connexion we may notice two. There is the sort of philosopher who scans the heavens for archetypal meanings; as Plato did for the meaning of 'Justice'. There is in contrast the sort of philosopher who is more interested in the ordinary meanings of words as used by ordinary people in the ordinary way. Both kinds of philosopher may be found among those who have interested themselves in the philosophy of mind. Brentano, for example, found the meaning of 'mental' in the scholastic conception of 'intentional inexistence' or 'reference to an object'. Prof. Gilbert Ryle finds with evident satisfaction that there is a sense of the word in which to be 'mental' is to be 'queer in the head'. Plato and Brentano contemplated a world in which there are aristocratic meanings. Ryle's world is a classless society of meanings.

The scientist's reaction to the diversity of meanings would seem to be different again. He is not especially interested in actual uses of a term or with archetypal meanings, but with uses actual or possible that can be turned to scientific account. Meanings can be distinguished as we distinguish, say, beetles, divided into orders, genera and species; or they can be considered as values of a variable. 'Good' is a word in point. It is clearly a word of very many meanings. This is not a matter of ordinary ambiguity. It is not just a matter of six or seven meanings but rather of six or seven hundred.

We can start, for example, with the simple degenerate but actual case in which 'This is good' means what is meant by 'This is what I like here and now'. In a judgement of this form there are three constants other than the thing or situation judged to be good. There is 'I' the agent, 'like' the psychological function that is exercised and the 'here and now' which defines the conditions under which the agent exercises the function in relation to the thing judged to be good. Replacing the constants by variables we obtain a formula of the form 'x exercises the function y under conditions z' in relation to the object upon which the judgement is passed. We may then assign any of a large number of values to each of the variables and obtain an indefinitely extensive series of possible meanings to the judgement of value. Of these a very considerable number may be of scientific interest and importance. Thus, substituting values for the *agent variable* we obtain judgements of the type 'This is what *I* should like, what *all members of my social class* would like, what *all members of my culture* would like, what *all men* would like, what *all sentient beings* would like. For the *function variable* we may substitute liking,

approving, needing and so forth. From variations of the *conditions variable* we obtain judgements of the form: This is what x would like under *all ordinary* conditions, or what x would like under an immense variety of other specifiable conditions.

There are possible, and interesting, theories of value which would take 'needs' as the constant for the *function* variable. Many possibilities are left open in regard to the *agent* and the *conditions* variable. The following, for example, are two interpretations of the meaning of 'This is good' derived from the basic formula in the way described:

(1) This is what would give satisfaction to the needs of any individual on the whole and in the long run.

(2) This is what would give satisfaction to the needs of any human social group on the whole and in the long run.

The first interpretation gives meaning to the statement that the good for the individual lies in the 'integration' of his personality. The second gives meaning to the statement that the good for a society consists in an analogous sense in the integrated satisfaction of social needs. On both interpretations the question whether this or that is or is not good becomes an issue of fact—to be decided by observations of actual needs under particular conditions and by inferences from these observations to the needs that would arise under other conditions.

The procedure here consists not in selecting from common usage one particular meaning but in suggesting or prescribing a meaning. This is common scientific practice, but there is nothing arbitrary in this. Meanings are defined and prescribed in accordance with principles. The term as defined (i) must have application, (ii) it must have application to 'a natural kind', and (iii) its applicability must be determinable in any given case by objective tests. Psychologists have not been quite so good in the use of these techniques as other natural scientists, but examples can be cited even from psychology.

V. Biological concepts of value

This can be usefully done for the slippery concept of value. It has already been done, in a sort of way, by the biologists in their discussion of biological needs, biological norms and in their treatment of 'survival values', but as commonly employed these are poverty stricken and basically unscientific concepts. They are poverty stricken in their limitation to the more elementary needs, for food, the avoidance of injury, reproduction, etc., and they are unscientific in that they do not correspond to any natural grouping of the facts. The expression 'biological needs' can in fact be defined only as denoting those needs in which, up to the present, biologists have happened to take an interest. The restriction arises from the accidents of history. There is no logic in a principle which includes the need for food, but excludes the need for change or variety of food, or if it includes these excludes the more specific food needs of the gourmet. There is no logic in a concept which allows the study of the food choices of rats in an experimental cafeteria but forbids the study of food choices of human beings in a self-service restaurant. There is no logic in a conception of biological needs which covers the case of a bird building a nest but excludes the case of a community building a cathedral. There is no logic in a concept which covers the interest of a cat in a mouse, but excludes the interest of the scientist in the cat's interest in the mouse.

The illogicality consists not, of course, in distinguishing primary from secondary needs,

original from derived needs,[1] but in linking these distinctions to the distinctions between the special sciences. The basic illogicality is the illogicality of a distinction between psychology and the other biological sciences.

We are led to the same conclusion when we look a little more closely at the traditional concept of 'survival value'. An adaptation is said to have survival value if through that adaptation life is prolonged. Man, by a suitable change in his metabolic habits might, maybe, prolong his life for a thousand years. Perhaps he could do so by hibernating like a tortoise in the winter and reducing the tempo of his waking life to that of a sophisticated cabbage. Would there be good sense in saying that such an adaptation had survival value? Obviously not. Survival value is at least a two dimensional variable. The life that any living thing endeavours to prolong is *life of a certain kind*. An adaptation can be said to have survival value for any creature when it either prolongs or increases the quality of the life that is characteristic of the creature in question. In general the quality of life characteristic of any creature can best be defined in psychological terms.

VI. Value versus fact

To redefine the evaluative concepts of biology in terms of 'psychological needs' reduces the gulf that divides these evaluative concepts from those of traditional philosophy and the gulf is reduced to the extent to which the good for man is defined in terms of the fulfilling of his needs. The concept of this good as that which would give satisfaction on the whole and in the long run satisfies some at least of the scientific requirements. It has application and its applicability is subject to observational tests.

The concept is of course subjective in the sense that it refers to 'needs', but is objective in the sense that what a man *would* desire under specified conditions, like what he *does* desire under specified conditions, is an issue of fact. In a sense 'good' is a matter of taste, but it is not a matter of taste to say what a man's tastes are or would be.

There remains an ultimate relativity. What is good for man may not be good for other creatures. The moralist has said that he would rather be a dissatisfied philosopher than a contented pig. The pig's comment has not been reported.

The pig no doubt would say that he would much rather be a discontented pig than any self-satisfied philosopher. Even the pig has its own distinctive good, needs which it would satisfy, if it could, on the whole and in the long run. Each kind of creature endeavours to persist in its own being, and in so persisting to enjoy its own distinctive quality of life.

So, too, perhaps with cultures, so too with human types. The colour blind may have little need for pictorial art, the tone deaf little need for music. The satisfactions of mathematics are not for the mentally defective. And in the last analysis so it may be for individuals. What is one man's meat may be another's poison. But this is not a matter for the throwing up of hands or the shrugging of shoulders. What is good for Peter may be different from what is good for Paul. The good for Peter may perhaps need to be tailored to his individuality. But what this good will be, like his individuality, will be a matter of fact. What a man *would* desire if he were consistent, if he were reasonable, if he took into account all the things that he desires and their compatibilities, what he

[1] The distinction between primary and derived needs is in part merely methodological. Primary needs are those which in a given explanatory system are *taken* to be such. Secondary needs are those what within that system are *shown* to be such. What is a primary need in one system (e.g. Allport's) may be secondary in another system (e.g. McDougall or Freud).

would desire if he took a long view, what he would desire if he took into account the fact that some satisfactions pall more rapidly than others, what he would desire if he had developed to maturity: all these and many other possibilities are matters of fact, grounded in his nature. We cannot of course directly observe what he would desire under conditions that do not in fact obtain as we observe what he actually desires under conditions obtaining, but statements concerning the former are derivative from empirical observations concerning the latter. In this sense values are objective. It is in this sense that psychology is a normative as well as a positive science. In short, normative science is a *kind* of positive science. There need be nothing disturbing in the fact that values are relative to a *given* nature, or relative to a level of maturation or to a level of culture. It is to be expected that new values will emerge, new goods discovered with experience and with the acquisition of new powers. Though 'goods' may be in principle predicted it well may be that no man can know for certain what he wants until he gets it, and that when he gets what he wants he will want something else.

There need accordingly be nothing disturbing in the fact that it is for each individual to discover for himself his own distinctive principle for the harmonization of his goals. To make this discovery calls for the exercise of his highest faculties of imagination, intelligence and reason. It is by reason, and the resources of science that this issue of fact will be decided—what is it that a man would desire on the whole and in the long run if he were consistent?

(*Manuscript received* 6 *May* 1953)

HYPOTHETICAL TERMS AND RELATIONS IN PSYCHOLOGICAL THEORIZING[1]

By W. M. O'NEIL

University of Sydney

I. *Three types of hypothesis* (pp. 211–213). II. *The hypotheses of Tolman, Spearman and Hull* (pp. 213–216). III. *The views of McCorquodale and Meehl* (pp. 216–218). IV. *The need to be clear* (pp. 218–220). *References* (p. 220).

I. Three types of hypothesis

Scientific theorizing, in so far as it differs from scientific observing, embraces two major activities. One of these is the making of guesses or hypotheses, and the other is the making of inferences. It aims first, at the explanation and prediction of observed phenomena; and secondly, at the systematization of our knowledge about these phenomena. The first aim is achieved when, subject to certain specifiable conditions, one is able to draw correct conclusions about the phenomena from other considerations. The second is achieved to the extent that one is able to draw conclusions about superficially diverse phenomena from a central and parsimonious set of considerations (premisses). Some of these considerations may be the direct product of observation, i.e. empirical generalizations. But ordinarily all the premisses needed for the required deduction are not available as observational knowledge. The theorist in that event must guess about what is the case.

The observed data of which an account is to be given commonly occur in psychology in one of at least two patterns. The first is more typical of the area of work to which the word experimental is applied; and the second is more typical of the psychometric area. In the first, there is some class of behaviour (the *dependent variable*) on the one hand, and on the other a number of observed conditions (the *independent variables*) deemed to be relevant to the former. In the second pattern, there is a number of behavioural performances, any one or more of which may be treated as the dependent variable; and commonly the other performances and the total set of instructions and tasks constitute the independent variables. In each situation, it is the dependent variable of which an account is to be given and the independent variables are used as premisses from which the former is to be deduced. However, as not all the terms of the required conclusion are to be found in the premisses made available by observation, additional propositions linking the conclusion to them must be provided. Thus, in one sense, the hypothesis introduced for this purpose may be said to be about what mediates or intervenes between the content of the observational premisses and the conclusion about the phenomenon. In the experimental context, Tolman (9) has spoken of the content of such hypotheses as *intervening variables*. Lewin (4), Hull (2), and others, have used the word *construct* in a broadly similar way. In the psychometric context, the word *factor* has been widely used. And Spence (7), in the course of making some distinctions about the material content of various theorists' hypotheses, has noted in passing some formal similarities between the experimentalists' and the factor-theorists' concepts. More recently, McCorquodale & Meehl (5) have distinguished on formal grounds between two kinds of assumption, which

[1] Based on a paper read to the Annual General Meeting of the British Psychological Society, Oxford, 1952.

they propose should be spoken of as intervening variables and hypothetical constructs respectively.

It is the primary intent of this paper to argue for the recognition of three formal types of hypothesis of this general sort, which for brevity will be called the *hypothetical relation*, the *uncharacterized hypothetical term* and the *characterized hypothetical term*. The word *term* is being used in the logical sense, as in the distinction between term (or quality) and relation, and not in the linguistic sense, as in the distinction between term (or name) and whatever the term names. Further, it is recognized that hypotheses are propositional in form.

The three types may be illustrated by some partly fanciful accounts of verbal association. The phenomena or dependent variables to be explained or predicted are: (i) the association of two or more words, i.e. for instance, the tendency for a person to say 'fork' in response to the request 'Say the first word that comes to mind when you hear me say "knife"', and (ii) the strength of the association, i.e. the frequency with which the response word is given to the stimulus word (other definitions might be preferred, but this will serve the present purpose).

First, let us take a quite radical empiricist account of these phenomena. It might be assumed that verbal associations result from the joint occurrence of the words in the person's past experience, and that the strength of the associations is a function, say a positive decelerating function, of the frequency of these past joint occurrences. No more has been assumed here than that the observed phenomena are *related* in certain ways to specified observed or readily observable antecedent conditions. All the laws of association, e.g. contiguity, similarity, contrast, simultaneity, recency, primacy and effect, when formulated in this radical empiricist way, are examples of *hypothetical relations*. So also is the law of gravity when it is taken to be a statement of the attraction between particles of matter as a *function* of the masses and distances of these particles.

Secondly, let us take the simplest additions to the above assumptions that might be made by someone not quite content with such positivistic formulations. He might ask that some rationale be provided for the conditioning of present associations by joint occurrences over and done with in the past. He might assume that in the past the joint occurrences made some alterations in the nervous tissue or in the mind-stuff (if he cares to believe in the second), alterations which persist into the present. These relatively lasting 'traces' bring about the verbal associations, and their strength is reflected in the strength of the associations. He would probably want to reformulate the laws of association to which he subscribed, by saying such things as 'the strength of the association trace is a positive decelerating function of the frequency of the joint occurrences'. Although he assumes an entity or state (i.e. a *term*) to help him bridge the gap between independent conditions and dependent phenomena, he has virtually nothing to say about the nature of this entity or state. Perhaps footprints in the sand suggested the notion of the 'trace' to him, or perhaps he might try to make clearer what he means by likening it to footprints—the more frequently the foot has stepped on the one spot, the deeper the impression. But he would insist that the 'trace' is not a footprint, nor indeed literally or pictorially like a footprint. It has roughly analogous laws, but no other likeness. And he does not say what a 'trace' literally or pictorially is like. Thus the term that he has assumed is almost entirely uncharacterized; all that is said about what it is, is the rather vague 'relatively persisting alteration of the nervous tissue' (or 'of the mind-stuff').

The history of psychology affords many examples of such *uncharacterized hypothetical terms*; faculties, abilities and factors conceived as powers of the mind, dispositions, determining tendencies and mental sets conceived as states of preparedness or of inclination of the mind, and instincts, wishes and tensions conceived as energies or sources of energy have been among the common instances. Although, in the formulations in mind at the moment, these have been regarded as entities, states or processes, we have been told singularly little about what they are and a good deal about how they work upon or are worked upon by other things.

Thirdly, let us look at what a more pictorially minded or a more speculative theorist might add to the second account. He might suggest, for instance, that the 'trace' consists of a lowered threshold of synaptic resistance in certain sensory-motor arcs, or of a neural circuit in the cerebral association areas, or of a relatively persisting mental image of which we are ordinarily unconscious. That is, he tells us what a 'trace' is, what it is literally like. Thus he has *characterized* the hypothetical term.

It will be noted that among the options provided in the second and third accounts were some distinctly 'mentalistic' formulations (changes in the mind-stuff and persisting mental images). Too often in recent discussions it has been taken that hypothetical constructs (or terms) must be physiological. Though current thought would not be receptive to assumptions about non-physiological 'substances', there seems to be no ground for regarding non-physiological 'processes' as impossible content for hypotheses of this type.

II. The hypotheses of Tolman, Spearman and Hull

Some discussion of concepts employed by contemporary theorists may help to sharpen the general points made above. First, it seems desirable to establish where Tolman's intervening variables fit into the proposed scheme. In several places (e.g. (10,11)) he has set out means of arriving at an objective or operational definition of the intervening variables used in his theorizing. He suggests that the conditions relevant to the explanation and prediction of a rat turning left rather than right at a choice-point in a maze are such environmental variables as maintenance schedule, the goal object and available stimuli, and such organismic variables as heredity, maturity and physiological state. To supplement the propositions about these antecedent conditions (so that appropriate inferences can be made), he assumes intervening variables such as *demand* which obtains between maintenance schedule and turning left. However, wishing to avoid a charge of arbitrariness in making these assumptions, he resorts to an operational definition of them. For instance, *demand* is to be defined by a series of experiments in which independent variables other than maintenance schedule are held at some standard value while it is systematically varied. *Demand*, he maintains, will then 'betray' itself in variations of the proportions of left-turnings at the choice-point. In the case of food deprivation within the framework of the normal maintenance schedule, Tolman indicates that the obtained data may on graphic exhibition be seen to reveal a growth and decline function.

Two interpretations of Tolman's intervening variable seem possible. First, and more consistent with the operational definition approach just outlined, it is the way in which the dependent behaviour (or some aspect of it) is affected by some specified condition(s). It is a functional relation assumed to hold between two or more observed variables, when

other relevant variables are having a constant effect. Secondly, and perhaps more consistent with his discussion elsewhere, it is a virtually unspecified something which is the medium through which the specified antecedent conditions have the specified effect upon dependent behaviour. In his early treatment(9), where he speaks of them as immanent determinants or the immediate causes of behaviour, it is apparent that the intervening variables are taken to be qualitative states or processes rather than functional relations. Later in his choice-point analysis a curious situation arises. He begins by stating that behaviour is some function (f_1) of the complex of environmental and organismic conditions. The intervening variables are then introduced piece-meal through the defining experiments outlined above, and the state of the analysis is schematically exhibited as follows:

Environmental variables	Intervening variables	Dependent behaviour
Maint. sched.	—f_2→*demand*	
Goal object	—f_2→*appetite*	
Stimuli present	—f_2→*discrimination*	—f_3→proportion of left turnings
Responses required	—f_2→*motor skills*	
.	—f_2→ .	

Tolman then admits that he has to face the solution of the f_3 function. Now if the dependent behaviour is a function of the set of intervening variables, they must surely be terms and not relations. Yet when they were deemed to 'betray' themselves through the defining experiments, what were exposed were relations and not qualitative. That is, if they are terms all we can say about them is how they are related to certain antecedent conditions and in prospect how they interact to produce the dependent behaviour.

The second interpretation is paralleled by the usual treatment of factors and traits by the factor-theorists. Spearman(8) seems to have recognized two independent parts of his Two-factor Theory. One, of which he was certain, was the presence of a factor *g* entering to varying degrees into all cognitive activities and of numerous *s* factors one of which was specific to each kind of activity. In this part of the theory, the only assertions about the factors concern their *relations* to observed performances and their multiplicative contribution to correlations. The other part, about which he was never more than tentative, consisted of such positive (as distinct from relative) assertions as that *g* was a central fund of cognitive energy or power, and that the *s* factors were engines or channels through which the energy was applied. The general factor in the former part of the theory is an uncharacterized hypothetical term; it is a something, we know not what, which is related in specified ways (the varying saturations) to all cognitive activities. In the latter part, it is characterized, admittedly rather vaguely, as a fund of energy. That the two are independent (i.e. one may be rejected and the other asserted or rejected without contradiction) is indicated by the fact that only the former has ever been tested in any explicit way, and only it is required to infer the observed correlations between cognitive performances. Had we some non-circular way of establishing *s* factor saturations, the former part would suffice for the inference of the actual scores. As it is, this cannot be done even upon resort to the second part of the theory.

Most factorists have left the hypothetical terms uncharacterized; Kelley is one other exception with his reference to genes. Often, they do appear to characterize their factors, but what in fact they do is to say how they are related to observed variables (the distinction between general, group and specific factors is made in this way) or to what observed variables they are related (the distinctions between cognitive and conative, or among

verbal, number, spatial and memory factors are made in this way). For instance, the definition of *verbal group factor* says merely that a hypothetical term is related in specified ways (high saturations) to all verbal performances of such and such sorts and in other ways (zero saturations) to all other sorts of performance; but it says nothing about what it is that 'saturates' some performances and not others. If I have not misunderstood it, the argument of Burt (1) that factors are principles of classification involves a rejection of the interpretation of factors as terms and is equivalent to their interpretation as means of stating a pattern of relations between observed performances.

An example of the characterized hypothetical term is to be found in Hull's notion of *drive* (2). Departures from optimum conditions for individual or species survival constitute states of need. These, often in the form of changes in the state of the blood stream, act as stimuli upon interoceptors and in conjunction with external stimuli evoke behaviour. Both Tolman's *demand* and Hull's *drive* bear upon what the former has called the 'persisting to and from' of behaviour. But Tolman's account is of the rise and fall of such persistings, whereas Hull's is of its being brought about. *Drive* is a positively defined agent (or term), i.e. certain qualities are predicated of it. Whereas *demand* is, as we have seen, either a functional relation between an environmental variation and a behavioural variation or an uncharacterized term standing in that same relation to the environmental variation. If there is such a state as *drive* we would, if we had the appropriate techniques for observing it, recognize it when we saw it. But granted these techniques, we would not recognize *demand*; whatever its properties may be, we would be cognizing them for the first time. The one we could recognize through its alleged distinctive qualities, the other we could identify only by its fruits.

When a characterized hypothetical term (or, in the usage suggested by McCorquodale & Meehl (5), a hypothetical construct) is introduced, it is necessary for prediction and explanation to introduce as well the relations into which it enters with observed terms. It is presumably such relations that are being referred to when McCorquodale & Meehl speak of the 'empirical law' associated with the hypothetical construct. Commonly, these relations are hypothetical too; although occasionally they have been empirically verified in other situations. A case of the latter is to be found in those instances where the perturbations in the orbits of known planets have been accounted for by the assumption of an additional but as yet unobserved planet (a hypothetical term). It is taken for granted that this hypothetical planet will work in accordance with the laws verified for observed planets. Hypothetical terms, of course, are not conceived as being totally unlike anything in the world of our experience; to that extent their assumption carries with it the assumption of a particular way of working, viz. that of the things they are like. Drive stimuli are deemed to be like external stimuli in some ways and to that extent are deemed to work like them. (Analogy is always an important source for the content of hypotheses). However, to make his assumed terms serve the purposes for which he introduced them, Hull has to assume other relations, such as the one linking drive stimuli, habit strength and reaction evocation potential, viz.

$$_{S}E_{R} = D_{S} \times {}_{S}H_{R}.$$

Incidentally, the distinctions being made here are quite independent of the issue of whether hypotheses are to be taken as guesses at what exists but has not yet been observed or as convenient fictions useful in ordering the otherwise bewildering world of

experience. In making them, the writer has probably revealed his leaning towards the former of these views; but that is also beside the main points being urged.

III. The views of McCorquodale & Meehl

It seems important at this juncture to relate the proposed triadic division to the twofold division of McCorquodale & Meehl. Clearly, what is here called a characterized hypothetical term is the same or almost the same as their hypothetical construct. And broadly, the other two types proposed here are a subdivision of their intervening variable. It is felt, however, that what they say about the latter is somewhat confused and cannot in all instances be supported. To bring out these points it will be necessary to examine in some detail their proposed bases of division.

They state that intervening variables 'merely abstract the empirical relationships' between the observed variables, that they are 'merely short-hand summarizations' of the empirical data, that the statement of them 'does not contain any words not reducible to the empirical laws' and that 'the validity of the empirical laws is both necessary and sufficient for the "correctness" of the statements' about them. They state, on the other hand, that hypothetical constructs 'involve the supposition of entities or processes not among the observed', that they 'have a cognitive, factual reference in addition to the data which constitute their support', that the statement of them 'involves words not wholly reducible to the words in the empirical laws' and that 'the validity of the empirical laws is not a sufficient condition for the truth' of the constructs.

It may be noted that the contentions made about hypothetical constructs can be applied to both of what were called above the two parts of Spearman's theory. Both parts involve the supposition of unobserved entities (virtually unspecified somethings in the one, and a fund of energy and several engines in the other). And if such things as the fulfilment of the tetrad criterion by tests having such and such content are the data which constitute their support, then both have cognitive, factual reference to something other than these data (admittedly in the first part, this hardly goes beyond 'there is a something'). The empirical laws in both are such relations as the factor saturations and the product equation, but the statement or definition of the factors is not wholly reducible to these laws (again, admittedly in the first part, there is very little left). And finally, even though the tetrad criterion were fulfilled (and that presumably would be regarded as establishing the 'validity' of Spearman's empirical laws), the truth of the factors would not be established—for one thing there are alternative factorial solutions and alternative interpretations of the 'two-factor' solution. That they do apply, if only weakly in certain respects, to what is here called an uncharacterized hypothetical term, may render doubtful the writer's identification of this as a subspecies of the McCorquodale & Meehl intervening variable. But if it can be successfully argued that at least some of them apply also to the hypothetical relation, and that one concept which those authors used to exemplify *intervening variables*, is indeed a term and not merely a relation, this doubt should be removed.

A mild objection will first be offered to the abstractive-constructive basis of differentiation. All types of hypothesis involve suppositions not among the observed. In many of the traditional discussions of scientific method, the imaginative efforts of scientists like Kepler have been underrated. It is said, for instance, that he merely abstracted from his data the laws of planetary motion, and that consequently these are merely 'empirical'

laws rather than those superior beings, 'rational' laws. But Kepler is reported to have formulated many hypotheses, each of which he tested on the data that he had. Was he adding something not among the observed in his earlier unsuccessful attempts, but merely taking out something already there in his final and successful attempt? If some epicyclic geocentric set of laws were formulated that fitted the data as well as Kepler's laws do, would they also be already among the observed and merely abstracted from amongst them? In at least one sense, we do not observe the ellipticity of the motion of the planets; we observe their positions in the heavens and note that the hypothesis of ellipticity is a good fit and that the hypothesis of circularity is a poorer fit. If we could get out where Archimedes wanted to stand in order to press on the end of his lever, we might be able to observe the ellipticity of planetary orbits. Now if the hypothesis is like Kepler's, an assertion of certain relations entered into by observed entities, states or processes, there is no difference between the 'empirical laws' and the hypothesis (assuming it to be empirically confirmed). If this is true, then the contentions about the reducibility of one to the other and about one being necessary and sufficient for the other can serve only as means of asserting this identity. It is another matter, however, if there is any suggestion that the observed data 'which constitute the support' of the laws are sufficient as well as necessary for the truth of the laws. Propositions which literally summarize the data are particulars, whereas the laws are at the very least the super-altern universals of these particulars.

Certain objections are to be offered also to McCorquodale & Meehl's analysis of Hull's concept *habit strength* ($_SH_R$), cited by them as an intervening variable. First, it is something other than the 'empirical laws'. Hull himself spoke of it as 'the perseverative after-effects of the no-longer-existent compound events represented by S, R, G, and N'(3) and treated it as something which interacted with present external stimuli and drive stimuli to produce *reaction evocation potential* ($_SE_R$). What it is precisely he does not say other than that it is some perseverative process in the organism (and from his general inclinations, one might presume that he considered the process to be going on in the nervous tissue); his statements about it are concerned principally with how it works, i.e. are statements of its laws. Thus it is something more than an empirical relation between observed variables; it is something which enters into the relations specified with observed and other assumed variables. Nor can it be agreed that if 'Hull's proposed "grand investigation" of the Perrin-Williams type should be carried out and the complex hyperspatial surface fitted adequately over a wide range of values, it would be meaningless to reject the concept "habit strength" and still admit the empirical findings'(5). The 'complex hyperspatial surface' is a hypothetical triadic relation between degree of deprivation, number of reinforcements and the number of unreinforced trials required for extinction of the response. That it is an adequate fit to the empirical findings of the sort indicated is confirmation of this hypothesis, which might consequently be regarded as an empirical law concerning the effect of the first two variables upon the third. But 'habit strength' is a further hypothesis about the medium through which the first two variables, now no longer acting, produce their effect upon the third which is occurring in the present. And it would be far from meaningless to accept the hypothesis that the effect is produced and yet to reject this hypothesis about the means of its production. It is true, of course, that the second hypothesis is so vague that we have no vast array of clear alternatives (beyond agreeing with the radical empiricists that we do not need

any hypothesis to cover the determination of the present by events over and done with in the past, that such things should be left by the scientists to those who have an irresistible urge to philosophize).

A final word may be said about necessity and sufficiency by making reference to the psychometric field. One may, for instance, accept a given correlation matrix as truly within the limits of sampling errors, the concomitant variation between the set of variables, and at the same time either accept or reject a 'general plus group plus specific factor' solution, even though this solution is a 'good fit'. One could not, of course, consistently accept both the matrix of correlations and the solution if the latter were not a 'good fit'. No matter how 'good' the fit, it would seem in factor analysis that other solutions may fit just as well. What we need to make the acceptance of the data and the rejection of the solution inconsistent is a unique solution that is a 'good fit', that is, a demonstration that only this solution implies the matrix of correlations. But even if we do accept the solution as an adequate statement of the functional relations holding between the observed variables, i.e. accept the empirical laws, we are not obliged to accept the identification of the factors with uncharacterized somethings in the organism.

IV. The need to be clear

Finally it may be asked why it is important to make any division at all among types of hypothesis other than in the interests of a certain logical neatness. The answer is in terms of how the various types of hypothesis are to be used if they are to fill the roles they were brought in to play. As McCorquodale & Meehl point out, a failure to make some formal distinction of this sort 'leads to fundamental confusions'. It would not be difficult to show that the periodic 'slides' made by Tolman between a relational and qualitative treatment of his intervening variables lead to difficulty and confusion.

Least can go amiss with the hypothetical relation, when it is identified as such. To make an empirical test of this type of hypothesis (and hence to use it for other than reckless explanation and spurious prediction), it must be, as Hull(2) put it, empirically anchored at both ends. This involves our having independently obtained observational data upon both the independent and the dependent variables and our having adequate means (definitions, criteria, etc.) of identifying these variables. We cannot test the hypothesis that the strength of a verbal association is a positive decelerating function of the frequency of prior joint occurrence of the words associated unless we have independently observed values of both the strength and the frequency. And, further, 'frequency of prior joint occurrence' is in need of further specification than was given in the discussion in section I. For instance, is one sustained joint occurrence to be treated as equal to one momentary joint occurrence? Or are immediately sequential occurrences in either order and strictly simultaneous occurrences to be treated equally? Without adequate definition of the terms which are alleged to stand in this relation to each other, we cannot know with any precision whether the case before us is one to which the law is supposed to apply or not. (These definitions would seem to be a matter for empirical inquiry rather than arbitrary fiat, but that is not especially pertinent to the present argument.) Deficiences in one or both of these two respects are not uncommon in less trivial psychological examples. Factor-theory, Lewinian field-theory and psycho-analytic theory are fertile sources. A factor example will be discussed briefly in a slightly more complicated situation below, and a psycho-analytic example is being discussed at length elsewhere. The broadest Lewinian instance can be found in the general equation $B=f(L)$, behaviour

is a function of the state of the life space. Ordinarily the Lewinian seems dependent upon a knowledge of the behaviour to be explained for any statement about the state of the life space.

Little is gained immediately with the uncharacterized hypothetical term as compared with the hypothetical relation in the task of explanation and prediction. Hull(3) does not fully answer the question he imagines a 'rock-ribbed positivist' like Woodrow putting to him on this point. The question, in essence, is 'If A is a function of xy, x a function of B and y a function of C, (where A, B and C are observed variables and x and y are hypothetical), why not use the single equation $A=f(B, C)$?' It is agreed with Hull, and with Tolman who seems to urge a similar point, that a convenient and possibly the only feasible way of beginning the attack on the problem of how several variables are jointly related to a further variable, is to take them one at a time or in small numbers at a time. Then having discovered how each one or each set of them affects the dependent variable when the effect of the remainder is constant, we may go on to discover their interaction in conditioning the dependent variable. But the first step in the above procedure will produce equations of the type $A=f(B)$ and not of the type $A=f(x)$. And although we can give the relations so expressed a name (such as *demand* in Tolman's case, or *reinforcement* as Hull does for the way need-reduction strengthens habit strength), there is no immediate gain but some long-term danger in imagining that these relations hide terms or are terms whose properties we cannot specify. This is needless reification. There may be other evidence, however, that suggests that in our statement of the relations so discovered we have jumped over the heads of some unobserved terms. Hull seems justified in urging that ${}_{S}H_{R}$ is not a relation directly between the independent variables he symbolizes as S, R, G and N on the one hand, and the dependent behaviour on the other. Various considerations prompt one to assume with him that there are some 'perseverative after-effects' of these variables standing, as it were, between them and the behaviour. Although the assumption of such a vaguely specified term furthers the explanation and prediction of the behaviour little if at all (about all that could be said is that it enables inference over the time gap by writing in a premiss with a time span in it), it has its value if it stimulates a search for the properties of this 'something'.

The same argument can be brought forward in respect of reified factors in the psychometric field. But apart from some vague suggestions that the factors might be genes, there have been few if any real clues for the guidance of a search for the properties of the factors. It is interesting, too, that in both the experimental and the psychometric contexts, these reifications have usually been given the names of what they affect, of what affects them or of the effects they are alleged to produce. This seems to be a manifestation of that besetting conviction that causes and their effects must be alike, i.e. that it requires a fat herdsman to drive fat swine.

There seems an acute temptation to regard the hypothetical term, whether it be characterized or not, as affording sufficient antecedent 'anchorage' in the use of the accompanying hypothetical relation. This is a temptation that is particularly attractive to factor-theorists, although experimentalists often succumb too. One of the modes of empirical verification of a given pattern of factor saturations is the deduction through the product equation (Spearman) or the sum-of-products equation (Thurstone, Burt and others) of the correlations between variables saturated in the specified ways. If the factors are regarded as terms, then they appear to be the anterior anchors required for the testing of the hypothetical relations specified by these equations and saturations.

But except in the original and literally two-factor theory of Spearman, there is no way of independently establishing the factors which are shared by those variables except their known correlation, i.e. what is to be deduced. Thus we have neither a test of the hypothetical relation nor an explanation by means of it.

In discussions of these types of hypothesis, little if anything is said about the fact that the hypothetical term is useless to us unless we have a statement of the relations in which it stands with other variables. It is independent of those relations and does not in any sense generate its laws. Thus ordinarily they must be assumed too (and tested in the ways already indicated). Sometimes, however, the term is known by observation in other situations and its laws may be brought over from those situations. This is so, even though the term has been borrowed on the slightest of analogies and with considerable modifications. Consequently there is some reduction in the 'degrees of freedom' in the hypothetical premisses being used as compared with hypothetical terms with hypothetical laws that have not been known to be applicable to those terms in other situations. And the fewer the independent components in a hypothetical set, the more readily is it tested. Where, however, there is no ground for the linkage of term and relation other than the convenience of the present situation, most of the tests of the hypotheses to which psychologists resort serve to confirm or disprove the hypothetical relations but leave the hypothetical terms untested; an example of this was given above in the discussion of Spearman's suggestions about the nature of *g*. If we want to make use of assumed terms we should be more concerned than we seem to be, to seek evidence on their positive properties; we should look for the equivalent of the 'vapour trails' of the physicists' particles or of the photographic signs of the biologists' genes. Hull appears to have been trying to guard against the difficulty of arbitrarily assorted terms and relations when he looked upon reflex-conditioning not as the unit of which all learning is composed but as the source of terms and relations which might be used in explaining other behaviour. At any rate, an extreme case of the independence of a term and its mode of operation is afforded by the entelechy hypothesis, a major bugaboo for Hull; what the entelechy is supposed to do is in no way bound by what it is supposed to be, and further it is like nothing that we know and behaves like nothing that we know other than what we do not understand and are trying to explain.

References

(1) Burt, C. (1940). *Factors of the Mind.* London: University of London Press.
(2) Hull, C. L. (1943). *Principles of Behavior.* New York: Appleton-Century.
(3) Hull, C. L. (1943). The problem of intervening variables in molar behavior theory. *Psychol. Rev.* L, 273–91.
(4) Lewin, K. (1935). *Principles of Topological Psychology.* New York: McGraw-Hill.
(5) McCorquodale, K. & Meehl, P. E. (1948). On a distinction between hypothetical constructs and intervening variables. *Psychol. Rev.* LV, 95–107.
(6) O'Neil, W. M. (1952). *Austr. J. Psychol.*
(7) Spence, K. W. (1944). The nature of theory construction in contemporary psychology. *Psychol. Rev.* LI, 47–68.
(8) Spearman, C. (1927). *The Abilities of Man.* London: Macmillan.
(9) Tolman, E. C. (1932). *Purposive Behavior of Animals and Men.* New York: Century.
(10) Tolman, E. C. (1936). *Proc. 25th Anniv. Celebr. Inaug. Grad. Stud.* Los Angeles: University Southern California Press.
(11) Tolman, E. C. (1938). The determiners of behavior at a choice point. *Psychol. Rev.* XLV, 1–41.

(*Manuscript received* 11 *June* 1952)

EDUCATIONAL MEASUREMENT: ITS SCOPE AND LIMITATIONS. A CRITIQUE

By W. KENNETH RICHMOND

Department of Education, Glasgow University

I. *Introduction* (pp. 221–222). II. *Dangers of overemphasis* (pp. 223–225). III. *Kinds of measurement* (pp. 225–228). IV. *Conclusion* (pp. 228–231). *References* (p. 231).

I. Introduction

'It is as well to remember that in matters affecting the human heart there is not the same reason that there is in the measurable sciences to assume that the most recently accepted view is the right one', writes Mr A. D. C. Peterson (6). While the vast majority of teachers and parents would almost certainly agree with this opinion, it may as well be conceded that a great deal, if not the bulk, of educational research is now conducted on the assumption that 'matters affecting the human heart' are just as determinable as any other natural phenomena. 'The sacred and hidden identity which no techniques can reach', as Maritain called it, is no longer held to be quite so inaccessible as it was. Cautious as the psychometrist and the mental-tester may be in announcing their conclusions, it is clear that their conception of measurement does not stop short at the purely physical. As Thurstone puts it: 'The faith of science that nature can be comprehended in terms of an order acknowledges no limitation whatever as regards classes of phenomena.' (10)

Late in the day as it is to protest that the nature of the presuppositions involved in educational and mental measurement remains largely unexamined, it has to be observed that the new movement raises a number of philosophical questions which cannot be ignored.

Certainly if we look at it simply from the point of view of educational research the change of emphasis during the past fifty years has been remarkable. A glance through the titles of post-graduate theses given in A. M. Blackwell's list, for example, reveals that whereas in 1918 subjects dealing with the theory and history of education were still popular, recent years have seen a heavy concentration of those which necessitate (or at any rate allow of) the application of experimental and statistical procedures.

Various reasons can be adduced for this ever-growing desire to bring education into line with the exact sciences. Among them, if not the first, was a widespread dissatisfaction with the vagueness and contradictoriness of educational theory as such. An age which was becoming increasingly unsure about absolute values and ultimate aims naturally took to an approach which promised some kind of definite answers rather than indefinite questionings. If the new study of education were to be recognized and to have an effective say in the modern world, it was felt that any conclusions which it reached must be capable of independent check. A typical expression of this view is to be found in Sir Cyril Burt's advice: 'It is impossible for the teacher to expect that his own impressionistic conclusions can win general acceptance unless he corroborates them by scientifically planned experiments, and expresses them in carefully defined scientific terms.' (2)

The passion for precision and the insistence on verifiable evidence, on which the modern school of educational thought prides itself, have, of course their roots in a movement as old as, if not older than, the Industrial Revolution. There are clear signs of its emergence when in 1864 the Rev. George Fisher of Greenwich Hospital School drew up a 'Scale Book', in which 'numbers assigned to each degree of proficiency in the various subjects of examination were recorded'. While the Victorians would have been the last to agree that human behaviour, like any other commodity, lent itself to quantitative assessment, it is evident that this modest beginning was inspired by a desire for something more objective and more reliable than an examiner's mark. Though Galton was not above assessing the efficacy of petitive prayer, however, the leading emphasis in quantitative assessment has come from America, and may be illustrated by Micheels & Barnes's(5) statement: 'measuring techniques will play an integral part in whatever new advances we are able to make, socially as well as technologically'.

Another root, with more subtle ramifications, is connected with the liking for experiment and observable things in preference to verbal discussions, a contrast which is reflected in Erasmus's dictum: 'Knowledge seems to be of two kinds, that of words and that of things.' Ever since Locke, and more particularly during the past hundred years, the educational thought of Western civilization has been penetrated through and through with the spirit of empiricism. The Platonic distinction between *episteme* and *doxa* has disappeared. It was no accident, therefore, that the nineteenth-century's recognition of the important part played by quantification in the natural sciences was extended to the field of educational theory, itself dominated by a philosophy of naturalism. 'I came to realise', confesses Dr Rice, 'that this (i.e. the claims of school men following different courses of study) was *all talk*, that no one really knew the facts because there were no standards to serve as guides. Then, one day, the idea flashed through my mind that the way to settle the question was to try it out.' From that day (in 1894) to this, research has spared no effort in its attempt to make comparative tests the basis for a systematic study of education.

While its original purpose may have been limited to estimates of educational *achievement*, from the start the advocates of the new method were nothing if not bold in their pretensions. Thanks to it, they claimed, hit-or-miss opinion would be converted into knowledge, guesswork supplanted by solid evidence. 'We have faith that whatever people now measure *crudely by mere descriptive words*...can be measured more precisely and conveniently if ingenuity and labour are set to the task.'(9) Thorndike's optimism, and his studied contempt of words, are typical of the spirit in which the work was undertaken.

To be fair, it has to be said that this initial aggressiveness is partly explained by the heavy opposition with which the educational measurers had at first to contend. It has to be said, too, that in so far as they relate to educational 'products' most of the claims have since been made good. Beyond question, measuring techniques have been of inestimable service in helping to bring about a more discriminate classification of pupils, more efficient assessment of attainments and improved methods of instruction. In Britain, it is true, it is a case of 'business as usual' where examinations are concerned, but even here that last stronghold of 'mere descriptive words', the essay paper, has not gone altogether unshaken.

II. Dangers of overemphasis

But to concede that measurement has its uses is not to agree unreservedly with its underlying faith. Fifty years ago it was the measurers who were the heretics. To-day the position is reversed. 'To talk sense is to talk in quantities.' Whitehead's saying is merely a modern version of Kepler's—'the mind of man is meant to consider quantity and it wanders in darkness when it leaves the realms of quantitative thought'—with this difference that to-day it meets with little or no opposition. So successful has been the attempt to professionalize the study of education that anyone who ventures to question the efficacy of experimental-statistical techniques is liable to find himself dismissed as a tiresome amateur. Too often, of course, he is. But while granting that the detailed techniques required for the carrying out of further research in any study are not the business of the average practitioner in the field, in the case of education there are many reasons why the distance between the two spheres should not be too great. Judging by the way things are going, it may be feared that the gap between 'educational research' and 'educational thought', not to mention the gap between research and general practice in the schools, is steadily widening and that in time it may result in an open breach. Without in any way seeking to discredit the practice of educational measurement, or disparage it as a vogue, we ought, nevertheless, to subject the theory on which it is founded to the most rigorous scrutiny.

Immediately we do this we are faced with some astonishing, not to say disturbing, discoveries. When they turn from their figures to any kind of verbal exposition the apologists of measurement indulge in some very curious writing indeed. So sure of themselves on the surface, the propositions with which they seek to underpin their elaborate structures, seem to be exceedingly shaky.

Let us take the first, and the one upon which in some sense all the others depend—Thorndike's 'Law' that 'Whatever exists at all exists in some amount'. The statement is clearly intended to be regarded as self-evident. To the present writer, on the contrary, it appears to be either meaningless or else pernicious nonsense. Even as an expression of faith in the future possibilities of measurement it is like saying that eventually the circle will be squared. Rembrandt's 'Night Watch' exists, but in what sense can it be said to exist 'in some amount'? To this the measurers' answer is that when the art critic speaks of Rembrandt as being a greater painter than Vermeer, Vermeer greater than de Hooch and so on, he is really instituting a comparison which amounts to an order-of-merit ranking and *ipso facto* measuring one against the other. But, surely, this answer is irrelevant. It ignores the necessary distinction between *assessment* and *appraisal*. Assessment is largely, or entirely, a quantitative affair. Appraisal, on the contrary, invariably involves a judgement of qualitative values. The one may be described as mathematical, capable of verification so far as it goes, but essentially impersonal. The other is critical. No useful purpose can be served by confusing the two functions: to do so is akin to saying that there is no difference between right and wrong in the arithmetical sense and right and wrong as the moralist or aesthete understands them. Moreover, this distinction holds 'even though for practical purposes we may try to correlate our estimate of quality with some scale of quantity. One picture is not two and three-quarter times as beautiful as another, nor is one crime three and a half times as heinous as another, even

if the price paid for two pictures, or the terms of imprisonment awarded for two crimes may be in these proportions.'[1] (8)

Which brings us to Thorndike's second proposition that, 'To know it thoroughly involves knowing its quantity as well as its quality'. Here confusion arises out of the double meaning attached to the word 'know'. Thorndike has failed to make the necessary and vital distinction between 'knowing that' and 'knowing how'. Take a simple illustration. If I am an analyst I may test a sample of brandy and discover that it is 80% proof spirit, forty years old, etc., with the result that I *know that* it is as good a brandy as any on the market. On the other hand, if I am a connoisseur, one who *knows how* its bouquet should affect nose and palate, I may come to the same conclusion, viz. that this is an excellent liquor. The two methods of testing, and the two ways of knowing, are quite different; and there is no point in pretending that they are the same. 'Knowing that' is a matter of fact finding, of quantitative assessment. 'Knowing how' is more in the nature of a personal and critical judgement. It is difficult to see how this distinction can be removed or how the most resolute attempts to convert qualitative into quantitative values can succeed.

Nothing deterred, however, the measurers advance their third article of faith: 'That the objective products produced rather than the inner condition of the person whence they spring are the proper point of attack for the measurer, at least in our day and generation'. Leaving aside the final saving clause, we must observe that the statement is another example of what Prof. Ryle has called 'the ghost in the machine' fallacy. It stems from a para-mechanical psychology which conceives itself to be a counterpart of Newtonian science. As Ryle (7) says: 'When the word psychology was coined, two hundred years ago, it was supposed that the two-worlds legend[2] was true. . . .As Newtonian scientists studied the phenomena of the one field, so there ought to be scientists studying the phenomena of the other field.' As an empirical study, psychology accordingly set itself to investigate, not the mind, but those segregated aspects of it which were henceforth labelled 'mental phenomena'.

Whether or not we agree that the two-worlds legend belongs to the past, it cannot be denied that educational psychology has been more or less exclusively analytical and that it has tended to forget that the whole man is not the same as the sum of the parts. Among other things, this has led to its becoming the perfect forcing ground for the theory and practice of measurement. Since every measurement represents a highly partial and abstract treatment of the 'product', no measurement or combination of measurements can pretend to give the whole story. The objection that the 'product' is not unique in anything like the sense that its producer may be said to be must count for something: nevertheless, there is wisdom in the Harvard Report's comment that 'to the extent that an object is truly unique and occurs only once it is not the stuff of science'. For all that, we are assured that: 'A complete measurement of, say, a composition might include an exact definition of its spelling, its usage of words, its usage of word forms, its wit, its good sense and so on and so on; and each of these might again be subdivided into a score or more of component elements.' (9) In the name of good sense, though, how does one arrive

[1] From time to time attempts have been made to record varying degrees of appreciation. Fechner, for example, tested the reactions of a number of subjects to different kinds of painting by observing changes in the rate of the pulse. It would be unkind, as it is unnecessary, to dwell upon the crudity of such experiments.

[2] I.e. the inward life of the mind and its 'outward' manifestation in behaviour.

at an exact definition of wit? And what would be the 'complete measurement', one wonders, of 'Mrs Battle's Opinions on Whist'?

This brings us nearer the crux of the matter. Clearly, any number of compositions (Lamb's included) can be graded by relative position, either by personal rating or by the consensus of opinion of those who are competent to judge, and from these positions it is possible to draw some kind of rough scale. But a rating scale is not necessarily the same as a ruler on which the units of measurement are clearly delineated. Moreover, no amount of mathematical ingenuity can ever overcome the objection that any numbers we care to assign to the products in question must always be arbitary. Supposing, for the sake of argument, that we have to assess a hundred essays, among which 'Mrs Battle's Opinions on Whist' happens to be one. If we adopt a conventional method of marking we may decide that Lamb's piece of writing, being perfect of its kind, earns 10 out of 10, whereas John Smith's illiterate scrawl deserves no more than 1 or 2. Alternatively, we may decide that there is no point in comparing them, since they belong to entirely different categories. The more mature the product the more fully does it reflect the identity of the person behind it, and the more necessary does it become for judgement to replace measurement. In the case of the ten-year old it may well be practicable to award numerical marks for spelling, vocabulary, sentence-formation and other mechanical skills, but it is significant that when it comes to assessing attainment in the higher stages examiners generally feel that justice is better served by thinking in terms of first, second and third *classes*. The distinguishing characteristics of a class are, first, that it is self-contained and, secondly, that it exhibits values which are peculiarly its own. To some extent these values may be considered as having both a quantitative and a qualitative reference, but normally it is the latter which decides them. The 'greatness' of Leonardo is not the same as the 'greatness' of Watteau. Each has his own dimension and there is no 'constant' to which either can be referred.

Now admittedly, the criticism that the numbers used in educational measurement are ordinal rather than cardinal has often been made, and it is not one which its exponents have taken at all seriously. With some justice, they point out that *all* numerical units of quantity are arbitrary and that just as measurement may be direct or indirect so it may be absolute or relative. Heat, for example, can only be measured by its effects and the fixing of zero on the thermometer to represent the point at which water freezes under certain conditions can, if necessary, be changed, as can the choice of units. But at least some kind of zero and some kind of unit can be agreed upon in cases like this. The trouble about relative measurement is that this is by no means easy: it is rather like having to use a ruler made from elastic instead of wood. Relative measurement cannot hope to tell us 'how much'. The question is whether its claim to tell us 'how much more or less than' in education can be substantiated.

III. Kinds of measurement

In an able review of the conditions under which measurement of any kind is possible, Mr C. D. Hardie (4) observes that there are three distinct classes of conditions which in turn give rise to three senses in which the use of the word 'measurement' can be held to be appropriate, according to whether the magnitude is intensive, fundamental or derived. In the case of intensive magnitudes the only condition necessary is that 'the property generates an order in those bodies that possess the property'. That is to say that while

it is perfectly possible to arrange 'bodies' of this sort in rank order, it is *not* possible to assign a quantitative value to the differences between them: and it makes no difference whether we call the series 1st, 2nd, 3rd, or *A*, *B*, *C*, since the unit of measurement which would enable us to do this cannot be found.[1]

With fundamental magnitudes, the conditions to be satisfied are more stringent. Not only must the property allow of serial ordering, it must also allow of some process akin to addition. In other words, 'the system *C* produced by this process from two bodies *A* and *B* both of which possess the property in question must be assigned a rank higher than the rank assigned to either *A* or *B* and the position of *C* in the series of bodies possessing the property must depend only on the position of *A* and *B* in the series and not on the order or method of their addition. When these conditions are satisfied it is customary to assign to the system *C* the sum of the numbers assigned to *A* and *B*. Weight and length are obvious examples of properties which are measurable in this sense.'

In so far as the physical sciences are concerned with direct measurement they are dealing with spatial and temporal (i.e. 'fundamental') magnitudes; but it does not follow that this is the only sense in which the use of the word 'measurement' may be said to be incontrovertible. For there is still a third category—the 'derived' magnitude—which the scientist has no hesitation in using and in terms of which educational measurement may be possible.

In this case the units (of which heat, density and electric voltage are common examples) do not admit of simple addition and can only be measured indirectly. They are established only on the basis of a clearly specified relationship which possession of the 'property' creates within a given series—a relationship which is governed by a numerical law.

Now the crucial question is 'To which of these three categories does educational measurement belong?'; and the answer must be that no one can say for certain, or at least that there is no general agreement. Clearly, any such measurement cannot be 'fundamental'; and there is something to be said for the view that if it is merely 'intensive' 'then it follows that most of the calculations which educationists have done are a waste of time'. There remains the sole possibility that such measurement is of 'derived' magnitudes. Unfortunately (or fortunately) human behavioural responses cannot be shown to answer to numerical laws, or to any mathematical constant, in the way that certain physical properties can. To be sure, Hardie is optimistic enough to believe that 'it is not likely to be long before some such laws are discovered, and I have little doubt that they will be found to justify much of the work that has already been done'; but the plain truth is that hitherto they have *not* been discovered—and that Hardie's optimism scarcely squares with his otherwise admirably ruthless logic. Whether or not 'derived' magnitudes are confined to the physical field remains a matter of pure conjecture. So long as it remains so there must be a considerable question mark attached to the whole business of educational measurement (apart, that is, from the measurement of educational 'products').

These objections apply most obviously, perhaps, to the calculation of test intelligence. It may be that the quality or property referred to as 'general ability' can be measured as a derived magnitude, analogous to the density of gases, or it may not. Spearman's 'g' may

[1] Pythagorean influences in mathematics have never ceased to be active, it seems. Almost at the beginning of the modern scientific movement the Frenchman, Viète, developed a method in which numbers were represented by letters.

represent a genuine entity, or it may, like the non-existent ether of the physicists, be nothing more than a symbol which has been found necessary and useful to support a hypothesis which has still to be worked out. Admittedly it looks as though the interval between an intelligence quotient of 110 and an intelligence quotient of 100 is something more than just a matter of placing on a scale; on the other hand, to say that the difference is one of 10 points tells us next to nothing, since no unit of intelligence exists. Nor have we any means of knowing whether the interval between 110 and 120 is the same as that between 100 and 110.[1] Certainly we cannot say that the exceptionally bright pupil with an intelligence quotient of 130 is twice as intelligent as the idiot.

Now that intelligence quotients are bandied about so freely by parents and teachers alike, the tendency to take the number as indicating an amount needs to be guarded against. If the present-day educationist has a besetting sin it is the folly of using test intelligence as a yardstick when, in fact, it is no better than a dipstick. A dipstick may tell us whether a tank is half or three parts full: by itself it cannot tell us anything more about the contents. Granted, it is better than nothing, yet we deceive ourselves if we suppose for a moment that in using it we are 'measuring' in the exact sense, for the numbers engraved on this particular dipstick may mean anything or nothing.

So far no mention had been made of yet another precondition of scientific measurement, viz. the necessity of isolating the property in question. Intuitively, and perhaps rightly, the layman has always sensed a certain absurdity in measuring something called intelligence without knowing what that something is, or at least how it is defined. Not altogether in vain the opponents of intelligence tests have protested that these tests do no more than test particular kinds of ability in a particular kind of situation and that any larger conclusions drawn from them need to be treated with reserve. The self-imposed rule of all test construction—to determine exactly what is to be measured in advance—has, they say, been disobeyed. In their view intelligence is a function of the whole personality from which it cannot be dissociated: in a word, it is not a separate 'essence' or entity. In this connexion, a glance at the past history of educational psychology helps to explain how it was that the quantitative idea of mental ability came to be such an *idée fixe*. Consciously or unconsciously, Wundt, Binet and other early experimenters inherited the late nineteenth-century explanation of personality as a bundle of apperception masses. In the days when the furniture of the mind *was* the mind, differential psychology had no great difficulty in isolating that aspect of it labelled 'intelligent'. Now that the psychology which prompted it is discredited, the attempted analysis is so infinitely more difficult as to become impossible—for, as Thurstone complains, any findings that accrue are always 'immersed in the irrelevancies of experience'. In the field of physical and biological science these 'irrelevancies' may be ignored but not, it seems, in the human field.

Consider for a moment just one of these 'irrelevancies', the influence of social class upon the calculation of test intelligence, first demonstrated by the researches of Decroly and Degand. That the most careful, individual type of test devised by Binet was not entirely 'objective' was strongly suggested by the fact that children from well-to-do

[1] It is true that once the mental age norms are fixed the calculation of some kind of 'annual increment' may be possible during the period when intelligence is still growing. But any such increment cannot be said to refer to intelligence as such any more than the newly opening petals of a flower can be said to 'measure' the growth of the plant as a whole.

homes scored over those who were deprived. Furthermore, when it came to applying these tests in more 'backward' countries it was found necessary to scale them down in difficulty in order that the results should remain statistically valid. This tinkering with the test material to suit an *a priori* principle according to which the distribution of intelligence everywhere follows the curve of chance is, to say the least of it, suspicious. In a recent work Prof. Zipf has demonstrated that the Gaussian distribution 'should not be expected whenever the quality measured in a collection of individuals is influenced by the fact that they are related parts of an organized whole'. In so far as the children tested are *schoolchildren*, then, it seems reasonable to argue that they are indeed 'related parts of an organized whole' and that to that extent alone the true distribution of intelligence may well follow a curve of its own.

A favourite argument adduced by mental testers in their self-defence is that which relates to the constancy of the intelligence quotient. It is easy to see, though, that it can be turned against them. Testing a performance is not the same as measuring a potential. For example, one test of an electric bell is whether it rings. If the current is below the minimum required strength no sound will be forthcoming: beyond that, however, the fact that the bell *does* ring is no indication of the strength of the current employed. The act of ringing remains constant, no matter whether the voltage be 6, 12, 20 or 50. Against this the defence may object that the analogy is false since the tests are so constructed as to render it impossible for any subject to 'burst the ceiling'—in other words, the potential cannot exceed the actual performance. This precaution, however, can only be taken by including more, and more difficult, items in the test and/or by imposing a time limit so as to rule out any possibility of the test being completed. As devices for ensuring that the quantitative assessment of intelligence shall have statistical significance, both these steps may be necessary, but from the point of view of qualitative thinking this rigging of the test situation is as false as it is artificial. By what prescriptive right is the child who solves three or four identical problems more intelligent than the child who solves only one? The question is not so fatuous as it may appear at first sight. It may be that, as with the bell, there is a great deal more behind the performance than can be deduced from the performance alone. To revert to the previous metaphor, the dipstick may have reached the bottom of the tank, but it cannot be said on that account that the contents are governed by a constant, for the horizontal dimensions are unknown.

IV. Conclusion

It is a truism, of course, to say that no more can be drawn from a test than what is put into it. More modest than they were, those who construct and administer tests are now ready to acknowledge that no test ever quite succeeds in measuring intelligence *qua* intelligence, though it may give a reasonably accurate index of certain kinds of ability. This pursuit of the *ignis fatuus* of 'pure' intelligence has led in turn to the intricacies of multiple-factor analysis, a region into which the student of education, if he ventures at all, soon finds himself bogged down in mathematical abstruseness. These factors, he learns, 'may be called by different names, such as "causes", "faculties" (!), "parameters", "functional unities", "abilities" or "independent measurements". The name for a factor depends on the context, on one's philosophical preferences and manner of speech, and on how much one knows about the domain to be investigated.'(10) No matter how we regard them, however, it appears that the existence of these hypothetical factors depends

on a concept of mind which is still predominantly mechanistic. If psychology has succeeded in getting rid of the idea of 'mind-stuff', it appears that it has not yet disabused itself of the idea of mind as a structure. In particular it clings to the belief that 'mental phenomena can be identified in terms of distinguishable functions, which do not all participate equally in every thing that mind does'. (10)

Now the multiple-factor theory cannot be condemned out of hand simply because it works on an assumption which cannot be proved. As the history of science shows, the very existence of a hypothesis often depends upon assumptions which are subsequently found to be incorrect and, therefore, abandoned. Physics and chemistry would never have reached their present stage had not earlier investigators explored blind alleys in search of phlogiston and the philosopher's stone. On the other hand, no one now supposes that attempts to square the circle, are other than misguided and better dropped. In the absence of any better guide, scientific theory has no choice but to work on its own assumptions. But can the assumption in question be justified? Is it even necessary? These questions are vital, for if it turns out that the concept of mind as structure is inadequate it follows that any theories based upon it cannot rightly claim to be scientific.

Is the mind to be thought of as a structure, or not? If Prof. Ryle is right this is an improper question: 'The "reduction" of the material world to mental states and processes, as well as the "reduction" of mental states and processes to physical states and processes presuppose the legitimacy of the disjunction "Either there exist minds or there exist bodies" (but not both). It would be like saying "Either she bought a left-hand and a right-hand glove or she bought a pair of gloves" (but not both).' (7)

To suggest, as we have done, that the vogue of educational measurement is not scientific may seem the height of temerity. Nevertheless, it deserves to be taken seriously. Of the pre-conditions of measurement in any sphere, at least one—the possibility of fixing a unit in terms of which the measurement may be made—remains unsatisfied: and when Sir Godfrey Thomson, seeking to by-pass this awkward fact, assures his readers that 'the idea of measurement has been so expanded during recent generations by the mathematical ideas of continuity, infinity, and limit that they (i.e. the pre-conditions) are becoming inadequate as a statement of the position'; (1) some of them may be forgiven for concluding that it is mathematics, not science, which is being brought in to redress the balance.

Science, it has been said, should not be used so parochially as to exclude theories unpatronized by the Royal Society. On the other hand, we can hardly allow Jepson's statement that 'scientific method is merely trained and organized common sense' to pass without qualification since the methods of logic, philosophy and the humanities have equal claims in that respect. Least of all can we accept the view that the application of this 'method' (whatever it entails) to the human problem will prove a panacea. Indeed, as the President of Harvard complains, 'Positive harm is done by claiming that the scientific method is going to save us. Indeed, something close to fraud is being perpetrated when this method is defined by implication as the process by which the physical and biological sciences have reached their present stage. I find quite inadequate the description of the scientific method given by most social scientists. To me it is not an accurate epitome of the way physics, chemistry and biology have developed; I doubt the wide validity of this alleged method.' (3)

Whether or not Dr Conant is right in wishing for a more restrictive use of the word

'scientific', there is no doubt that the methods of educational psychology partake of a somewhat hybrid character, existing as they do on the dubious borderline between science and philosophy.

Since that day in 1894 when Dr Rice was seized with the conviction that the thing to do was 'to try it out', the path of experimental psychology has been paved with good intentions. Varied and numerous as the experiments have been, however, they do not seem to have got very far by way of converting 'opinion' into 'knowledge'. Papers with such ambitious titles as 'The Measurement of Values' are still written, but on studying them, it is disappointing to find that they are mainly about such topics as the evaluation of handwriting or the forecasting of popular choice in restaurant menus. Great play is made of Weber's 'Law' and Fechner's 'Law' in the investigation of psychical phenomena, but in a context in which 'psychical' takes on the lowliest of meanings. Even when it is a matter of measuring degrees of sensation intensity it appears that an ounce of experimental method calls for a ton of mathematical processing by way of interpretation!

It would be unbecoming, no doubt, to deride these studies which are as yet, after all, in their infancy, for their imperfections and shortcomings or for not producing spectacular 'results'. We do not need Aristotle to remind us that 'to seek for utility everywhere is entirely unsuited to men that are great-souled and free'. Franklin flying his kite in a thunderstorm and Mendel pottering about with sweet-peas were not engaged in activities which, *at the time*, were conspicuously useful, yet both made discoveries of the highest significance. In any case, no one denies that the benefits resulting from educational measurement to date are not inconsiderable. There is the danger, however, that unless its pretensions are checked, an unconsidered and ill-informed acceptance of its postulates may vitiate our educational thinking as a whole. If, as seems likely, the statistical cult leads to a new scholasticism (and future ages may find it just as ludicrous to discuss what is meant by 10 'points' of intelligence quotient as to argue about how many angels can arrange themselves on the point of a pin), it is perhaps no great matter, though it may be timely to ask whether all students are profitably engaged in acquiring techniques of this kind.

Much of the foregoing may seem like a tilting at windmills. If only as a plea for more qualitative thinking, nevertheless, it may not be out of place; for while at the moment it may be true that the all-or-nothing proponents of educational measurement are in a somewhat chastened frame of mind there is no saying when their claims will be renewed or to what strange lengths they may go. It is in their favour that they live in an age in which the spirit of scientific humanism is still very much in the ascendant. Unquestionably, the quantitative thinker, with his realistic approach and analytical skill, will always have his place; but the crying need of the mid-twentieth century is not so much for fact-finding as for greater philosophical insight in *knowing what to do about the facts*. The trouble about modern science is that it has proved highly successful in extending man's control over nature without adding to his understanding of nature. Instrumentalism is not enough. An educational psychology which is preoccupied with the *disjecta membra* of mental life may well be in keeping with the spirit of the Age of Uncertainty, but to the extent that it is, its findings must be regarded as partial and inconclusive, no matter how impressively backed by empirical proof they may be. The climate of opinion being what it is, most of us must doff our caps to those who use a formidable statistical jargon, but assuredly the day will come when, once again, it will be acknowledged that the language

of 'mere descriptive words' is capable of refinements which are not to be despised and appraisals which no predictional arithmetic can compass.

References

(1) Brown, W. & Thomson, Godfrey (1921). *The Essentials of Mental Measurement*, p. 1 (footnote). Cambridge University Press.
(2) Burt, C. (1950). *The Backward Child*, p. 684. Third edition. University of London Press.
(3) Conant, J. B. (1948). *Education in a Divided World*, pp. 121, 122. Harvard University Press.
(4) Hardie, C. D. (1942). *Truth and Fallacy in Educational Theory*, pp. 128 *sqq*. Cambridge University Press.
(5) Micheels, W. J. & Barnes, M. Ray (1950). *Measuring Educational Achievement*, p. 18. McGraw-Hill Series in Education.
(6) Peterson, A. D. C. (1952). *A Hundred Years of Education*, p. 100. London: Duckworth.
(7) Ryle, G. (1949). *The Concept of Mind*, pp. 319, 22–3. London: Hutchinson's University Library.
(8) Streeter, B. H. (1927). *Reality*, p. 26. New York: Macmillan.
(9) Thorndike, E. L. *The Nature, Purposes and General Methods of Measurements of Educational Products*, pp. 16, 17. National Society for the Study of Education: Seventeenth Year Book.
(10) Thurstone, L. L. (1947). *Multiple Factor Analysis*, pp. 53, 55–6, 57. University of Chicago Press.

(*Manuscript received* 7 *November* 1952)

THE DYNAMICS OF LEADERSHIP IN ADOLESCENT SCHOOL GROUPS[1]

By L. W. SHEARS

University of London, Institute of Education

I. *Introduction* (p. 232). II. *Classification of the observed behaviour* (pp. 232–234). III. *Three hypotheses about general functioning* (pp. 234–236). IV. *Rigidity in group structure* (pp. 236–238). V. *Essential determinants in the emergence of leaders* (pp. 238–240). VI. *The development of leadership differentiation during adolescence* (pp. 240–241). VII. *The effect of revealed differences between the groups on the emergence of leaders* (pp. 241–242). VIII. *Summary* (p. 242).

I. Introduction

In this investigation, the aim was to examine normal school groups, with a history of continued membership, in situations in which they were so placed as to require organization to reach their goal. It was expected that in these situations the members would find it necessary to organize themselves in order to reach their goal and that in the course of this organization leaders would emerge. It was hoped to observe the general functioning of the groups, examine the formation of group structure, and the extent of its rigidity and note the conditions determining the emergence of leaders. Hence we could see whether adolescents are able to distinguish the kind of leadership needed in a particular situation and choose a person or persons capable of playing the role demanded.

The groups were chosen from two co-educational schools, one a non-residential county grammar school and the other a residential private secondary school, both preparing children for the General Certificate of Education. A junior group of 17 girls and 15 boys of average age 13 years 10 months and a senior group of 16 girls and 16 boys of average age 15 years 10 months were selected from the grammar school. A junior group of 20 girls and 16 boys of average age 13 years 7 months and a senior group of 24 girls and 17 boys of average age 15 years 11 months were obtained from the private school. The groups from the two schools were, therefore, approximately equal in size, age, and sex distribution.

II. Classification of the observed behaviour

The main technique employed was the observation of the groups in action. Each of the four groups was observed in a learning and in a social situation. For the learning situation the groups had to conduct their own English lessons over a period of three weeks and for the social situation they had to organize a dance or party for the other forms in the school.

In order to develop an observation sheet for recording behaviour as it was observed, a pilot study was made with another form at the grammar school. It was found that the action could best be described in terms of positive, neutral or negative contributions to

[1] This article reports certain aspects of an investigation which was presented to the Institute of Education, University of London, in partial fulfilment of the requirements of the Ph.D. Degree and which was made possible by grants from the Imperial Relations Trust, the John and Eric Smyth Trust of the University of Melbourne, and the Education Department of Victoria, Australia.

the advancement of the group to its particular goal. The positive aspects could be divided into leading and supporting, and the negative aspects into withdrawing and destroying, while the neutral position could be called acquiescing. It was found necessary to use these five categories of action in describing the behaviour of the members of the groups; but to do this it was also necessary to define them in terms of observable actions and their apparent functions. An understanding of the list is not possible unless it is realized that the isolated item must be considered in the context of its occurrence and relies for classification in certain instances on the interpretation of the observer. The functional definitions made were as follows:

Leading

'Leading' is used to describe those activities of a member of the group which are accepted by the group as satisfactory for the attainment of the particular goal and are acted upon. The person leading at any particular time may be called at that time, the leader. His activities may include:

Suggesting a certain line of procedure to the group which is considered and acted upon.

Giving orders, making statements, explaining difficulties, or asking questions which the group agrees to answer or carry out or accept.

Endeavouring to find out a person's possible contribution to group progress.

Asking questions of, or disciplining members of the group in an endeavour to get all members contributing to the attainment of the particular goal.

Making the initial suggestion as to who would be the best person to lead.

Taking over from the person leading and calling on a member to act in a certain way.

Making a final decision on a point which the group accepts.

Supporting

'Supporting' indicates activities giving external evidence of a member's active support for the leader at that particular time. It is illustrated thus:

Supporting a member who has already been suggested for the position of leader.

Carrying out willingly certain instructions of the leader such as giving out materials or reading on the leader's direction.

Offering his services to the leader. Raising hand to volunteer or to offer an answer to a leader's question. Commenting on the work in hand to the advancement of the group but acknowledging the authority of the leader. (In criticism of this kind, the tone of voice and the manner of speaking must be considered as important as the words spoken, in deciding whether the comments are supporting or destroying).

Disciplining, on behalf of the leader, members engaged in destroying activity.

Acquiescing

'Acquiescing' indicates a member's acceptance of the situation by activities which neither add to nor detract from action leading to the attainment of the group goal. For example:

Accepting a procedure, the situation or any leader without comment.

Carrying out an order given to the group by the leader.

Withdrawing

'Withdrawing' indicates those activities of a member that give evidence of his withdrawal from the situation at that particular time. For example:

Comments between members which have no significance for group progress.

Touching, tickling, grinning, giggling between members which does not interfere with other members of the group. Members working individually or in subgroups with no link between themselves and the rest of the group.

Criticism of the leader without endeavours to influence other members, but with consequent withdrawal of the person. Moving aimlessly about the room, taking no notice of group activity or sitting quietly but obviously not participating. Interest in the activity of a member of a subgroup who is withdrawing or destroying.

Destroying

'Destroying' describes those activities of a member that destroy the attempts by the group or the leader to reach the particular group goal. For example:

Refusing without attempt at excuse to carry out the directions of the leader.

Making comments or moving about the room individually or in subgroups in such a manner as to cause the leader to stop group progress in order to bring the person or persons to order.

Objecting to a person being leader whom the group has accepted.

Objecting to the accepted leader endeavouring to discipline another member who is withdrawing or destroying.

Correcting a mistake of the leader in a nasty or disparaging manner.

Disagreement with the course of action planned by the leader and endorsed by the rest of the group, expressed in such a manner as to halt group progress.

Refusal to take steps to keep up with group progress which causes delay and obstruction to the leader and group.

Criticism of the leader's activities with an endeavour to influence other members of the group to withdraw or indulge in destroying activity.

Remarks by the leader made in such a fashion as to arouse the antagonism of the majority of the members.

With these categories of action clearly defined, it was possible to have a consistent approach to the behaviour shown in the situations. It was also feasible to concentrate on the categories of leading and supporting on the one hand and withdrawing and destroying on the other, as the relevant behaviour tended to come in these categories. At one stage, there might be a development in leading or leading and supporting, then steady progress with occasional destroying and then another development in leading and supporting. Only rarely was the observer faced with the problem of all categories being active at once over any length of time. This fact, together with the simple code of symbols and an observation record sheet, enabled the behaviour of the groups in terms of the emergence of leadership in a particular situation to be recorded.

In order to obtain a measure of the members' participation in the situations, the running record of the behaviour was analysed and each member's activity for each minute described in terms of the five categories of action. The total number of minutes spent in each category for each individual during each situation was ascertained and these figures were expressed as a percentage of his total activity. As has been shown, the categories range from leading as the most positive to destroying as the most negative aspect of participation. The weights attached to the categories were: leading $+2$, supporting $+1$, acquiescing 0, withdrawing -1, and destroying -2. The weighted figures were added, taking notice of the signs and in this way a Learning Participation Score and a Social Participation Score was obtained for each member.

III. Three hypotheses about general functioning

Observation of the groups was directed initially by three main hypotheses. The first was that each group would tend to choose one or more of its members to take the lead. In the eight situations observed, this hypothesis was confirmed because leaders emerged in all the situations. Table 1 indicates the participation scores of the leaders in the various

situations and gives the next highest participation scores of the remainder of the group in each situation. It shows that there was a sharp fall in the scores at a certain point, supporting the hypothesis that the members above this point participated to a considerably greater extent than others in guiding the group to its goal. These were identified as leaders.

The second hypothesis was that the general functioning of the group would be more adequately described in terms of leading, supporting, acquiescing, withdrawing and destroying (as already defined) than in terms of the more usual dichotomy of leading and following. In each situation the groups were given a goal and left strictly to their own devices. In all cases action began and it proved possible to classify most of the salient behaviour under the five headings suggested. In contrast, the leader/follower classification, while it covered the few instances when everyone felt organized to their own satisfaction by one leader, could not do justice to the tensions, frustrations, rivalries of cliques,

Table 1. *Members of the groups who were leaders in the experimental situations observed, their participation scores and the highest participation scores of other members of the groups*

Learning situation. Participation scores				Social situation. Participation scores			
Grammar school		Private school		Grammar school		Private school	
Senior group							
Boy leader T.	138	Boy leader R.	194	Boy leader W.	118	Girl leader C.	163
Girl leader M.	56	Girl leader B.	180	Girl leader J.	101	Girl leader A.	137
Boy leader R.	52			Girl leader MA.	98		
				Boy leader WE.	98		
				Boy leader WR.	80		
Next highest participation score of any other member:							
	38		40		64		82
Junior group							
Girl leader M.	114	Girl leader L.	137	Girl leader V.	113	Girl leader L.	139
Girl leader S.	79	Girl leader A.	93	Girl leader M.	96	Boy leader D.	116
Girl leader V.	66	Boy leader D.	91	Girl leader S.	66	Girl leader A.	60
Next highest participation score of any other member:							
	29		60		34		47

individuals' obstinate refusals to participate, and withdrawals for various reasons which characterized the goal pursuit of these adolescent school groups. It seems likely that in varying degrees they characterize all groups of a size and duration sufficient to allow group relations and norms of behaviour to develop so that a classification with more than two variables will always be required.

The third initial hypothesis was that within the groups studied there would be a tendency for each member of the group to become identified in the minds of the other members with one of the categories of action described, subject to the qualifications implied in other hypotheses relating to the rigidity of the developed group structure.[1] The active striving of these groups towards their goal can be seen as a sustained attempt to achieve stability, to establish what can be called an orderly progress. One aspect of this attempt is that members of the groups tend to become identified in the minds of the other members of the group with one of the categories of action.

[1] See section IV, p. 237.

To check this hypothesis each member of each group was asked, both before and after the situations, which of the others they considered the most and the least suited to lead the situation in question. Those with the six highest and four lowest participation scores were then taken out and a study was made of their relative positions in the eyes of other members (*a*) before and (*b*) after the situation. In all cases the relation between rank and position and participation score either remained the same or became closer on the second rating. This might reflect merely an increased understanding of the ranking process, but it also suggests that some members tended to become identified in the minds of others with at least the more marked positive and negative aspects of participation. Further support for this view is found in the fact that some members included (either at the top or bottom) in the first ranking were omitted after the situation and others, omitted altogether in the first ranking but who had obtained either high or low participation scores, were included in the second ranking. In short, the members' choice after the situation was influenced by the nature and degree of participation shown by members with high or low scores, and that those members aroused either positive or negative expectations in the minds of the choosers.

IV. Rigidity in group structure

While it is likely that there is a general tendency for members of the groups to become identified in the minds of the other members with positive or negative aspects of participation the extent to which they remain so identified depends upon at least two factors relating to the rigidity of group structure. The first of these is the extent to which members feel that the urgency of attaining the particular goal has not altered. This statement can be illustrated by reference to the record of observation. For example, in the junior group private school social situation, one girl had alternated between destroying and supporting activity according as she failed to gain or gained her own way. As the following extract indicates, she had become identified in the minds of most of the group with the negative aspects of participation:

Session 1. *Part of minute* 41

Girl leader L. 'What will we do with the troublesome ones?'
Boy leader D. 'So and so is a problem. She wanted to be on the committee.'
All agreed.

However, as the arrangements for the cooking proceeded, the need for some variety in the menu was expressed, and the group was held up over this question. The girl in question offered to get her father to send down ice creams, and went out to ring up. On her return she went up to girl leader L.

Session 2. *Minute* 33

Girl leader L. (to So and So): 'Go on!' 'Give it!'
Another girl (to So and So): 'You give it out.'
The girl was put on a desk and announced that her father was sending down 60 ice creams for the party.
All cheered and clapped.

Minute 34

Girl leader L.: 'Three cheers for So and So.'

These were given and the idea discussed in subgroups. The problem of providing food had become an urgent one. So and So supplied the wherewithal to overcome the problem and the attitude of the members of the group changed towards her. This is further indicated in the members' rating which ranked her eighteenth before the situation and tenth after it.

A second example occurs in the social situation at the Senior Private School. The announcement to this group allowed one week to organize the social, and in the first session, progress was slow, with the leaders having difficulty in making progress.

Session 1. *Minute* 40

Boy, and girl leader C. shouted to get order. Another very rowdy.
Boy supporter (to rowdy one): 'Shut up.'
Girl leader C. (to rowdy one): 'For goodness sake, shut up or get out!'

Minute 41

Girl leader C. (to others out the front): 'We don't know what they want.'
Boy: 'Why not a mixture?'
Rowdy Boy again a source of trouble with three other boys.
Very noisy and no progress.

However, two girls suggested a Hallowe'en party in two nights' time instead of the suggested time. Permission was obtained and the whole scheme became a race against time; the urgency was increased, and the group presented a different picture in the sessions which followed. Girls C. and A. became the leaders, and their orders were followed implicitly by the rest of the group.

Session 3. *Minute* 3

Some clamoured around girl leader C. for their jobs.
Girl leader C.: 'I can't tell everyone at once.'
Girl Leader C. called for order. Group quietened. She began to allot jobs.

Minute 17

Girl leader C.: 'Now those with jobs may go and get moving.'
Girl leader A.: 'Remember, there's not much time.'
The group broke up and went out talking and hurriedly.

Here is an example of the effect increased urgency had on the consolidation of the leaders' position. Others who had shared the leadership during Session 1 became supporters with the rest of the group, and girl leaders C. and A. were identified strongly under the need for urgency as leaders in the situation. Why they were so identified and not some other persons will be discussed later.

It is suggested that the introduction of urgency into a situation can affect the structure of the group. It has been shown in these two examples that an individual's position in the hierarchy of the group in a particular situation can be altered either for a short time, as in the first example, or lead to increased leader status for some members and reduction of the status of others, as in the second example.

The second factor which may affect the stability of group structure is that, as the situation progresses, a leader is required to possess certain knowledge to enable him to continue to lead the group to the attainment of the particular goal. The stability of group structure is partly a function of the consistency of the leader in regard to the possession of this knowledge. If he does not possess the required knowledge at a specific time,

the structure alters to enable another member with the knowledge to lead the group at that particular time. The permanence of the alteration in the group structure is dependent on other factors which may or may not result in the restoration of the original leader when the group has overcome the present obstacle. Examples of alteration in group structure of this type can also be given from the records of observation.

(1) *Junior group grammar school—learning situation*

Session 2. *Minute* 18

(Previously girl leader S. had conducted reading, but on its completion was uncertain as to procedure.) Girl leader S. asked Boy to give his talk.

Boy: 'Haven't got it prepared.' Girl leader S. then asked another boy.

He replied: 'Left it at home.' Girl leader S. then asked a girl.

Girl: 'So have I.'

Girl leader S. (in desperation): 'Can anyone give a speech?' General confusion. Girl leader V. volunteered and went to the front.

Minute 19

Girl Leader V.: 'Quiet! I'm going to talk about horses.'

Boy (to girl leader S.): 'You should sit down.'

Girl Leader S. did not move. Group quiet as Girl Leader V. spoke.

Girl Leader V. continued to speak until minute 28, then stopped and answered questions. When the questions were completed, girl leader S. took control to allot the work for the next day. In this example, girl leader S. had reached a point where she did not know how to proceed. Girl leader V. stepped in and took over the lead, disciplining, giving a talk and then answering questions. In this case, the lead reverted to Girl leader S., but in another case, when one of the boys did not know how to handle questions on Macbeth, which the group were reading, girl leader V. took over the questioning and she retained the leadership throughout the session.

(2) *Senior group grammar school—social situation*

Session 2. *Minute* 27

Boy leader W. had been outlining the arrangements, but was pressed for information regarding the programme of which he was unaware. General confusion resulted.

Boy leader W.E. came out to the front. 'Will everybody please be quiet and sit down.' The group settled down.

Boy leader W.E.: 'At the moment these are the games.' All listened.

In this case, boy leader W.E. retained the lead until the end of the session, explaining the games and answering questions.

These examples indicate that the rigidity of the structure of these groups in the observed situations was reduced not only, as previously shown, by a change in the urgency of attaining the particular goal, but also by any inability shown by the leader in meeting the demands of the members of the group in regard to the attainment of the particular group goal.

V. Essential determinants in the emergence of leaders

To obtain other information which it was felt might be relevant to the emergence of leaders in these situations an interview was conducted. As the interview was to be used to collect information on friendships, workmate preferences for learning, social, and

general school situations, family relationships, home economic status and cultural level, personal, cultural and social interests, social functions experiences, responsibility position experience, social contacts with the opposite sex, and personal health, it was decided to use the guided technique, asking the factual questions at speed while allowing the subjects all the time they required on the very personal questions. To guide the interview (which lasted approximately one hour) a list of questions was compiled and these were asked in the same order for each individual. No attempt was made during the interview to assess personality or examine attitudes. Replies were rated afterwards in a standard fashion so that some comparative assessment of each individual could be made under the headings listed above. In addition, members were asked to rate their fellows on ability at English, ability at organizing, and on popularity, and each member was given a battery of tests relating to personality adjustment. These various scores and ratings were correlated with the learning and social situation participation scores. It is not possible to produce the large table of correlations, but the results obtained can be outlined.

In the senior groups scores which correlated with the learning participation scores at the 1% level of significance were, first, a 'skills' score based on the marks of the last English examination and, second, a score relating to group relationships which might be called a 'workmate preference' score. This was the score resulting from a preference test on the criteria of working with members as leaders of English lessons. For the social situation there was almost complete agreement between the participation scores and the other evidence. Correlating significantly with the participation scores were, first, the three scores relating to the skills and experience required for the situation (social interests, social functions experience and social contacts with the opposite sex), and secondly, the four scores relating to prestige for this situation and group relations (rating on organizing ability, popularity rating, social workmates preference score and pre-social situation scale score). In short, the evidence from the senior groups suggests that the relative importance of the main conditions for emerging as a leader is as follows:

(*a*) The tools required for the situation as measured by tests of skills and experience.
(*b*) The group relationships developed within each group.
(*c*) The present prestige in the group of the member or members based on activity which appeared to the members as being of a similar kind.

When the junior groups are considered, the position is complicated, in regard to the various prestige ratings, by the inability of the members at this level to identify the people who in fact possess the required characteristics. At this level all the prestige ratings correlate significantly with both social and learning participation scores, and also nearly all the group relations tests of preference based on the 'working with' criteria. However, the test scores which were not based on members' opinions suggested that for these younger groups also there were required tools. In the learning situation at the private school the marks in the last English examination correlated significantly with the participation scores. In the social situation, social functions experience and responsibility experience at the grammar school and, in addition, social interests and social contacts with the opposite sex at the private schools, gave significant correlations. It would appear that in the junior groups there were skills and experience in one situation which gave members prestige and this carried over strongly into other situations and enhanced the acquisition of further skills, experience and so on. The following deter-

minants, listed in order of importance, are suggested by the correlations as being related to the emergence of leaders at the junior level:

(*a*) The present prestige in the group of certain members arising out of other group activities.

(*b*) The group relationships developed within each group.

(*c*) The 'required tools' for the situation as measured by tests of skills and experience.

On the negative side, there is little evidence at either age level to suggest that the following determinants are related to the emergence of leaders in these experimental situations:

(*a*) The degree of personality adjustment of the members of the groups.

(*b*) Personal factors such as age, intelligence or health.

(*c*) Factors relating to the home environment of the member.

VI. The development of leadership differentiation during adolescence

Examination of Table 1 reveals that at the senior level the leaders were different in the two situations, while at the junior level the leaders were the same in both situations. This fact led to the hypothesis that the capacity to differentiate leadership requirements and to recognize the persons who in fact possess them develops during adolescence. This hypothesis was tested, first, by examining the ability of adolescents at the two age levels to rate their fellow members in regard to the possession of certain requirements. Seven-point graphic rating scales were devised for ability at English, ability at organizing, and popularity and, in addition, a list of seven statements, ranging from complete acceptance of a person as leader, through indifference, to complete rejection was given to each member in each situation, together with a list of all the members of the group. In this way ratings on ability to lead in a social situation and in a learning situation were obtained from the members of the groups. The ratings obtained on these five criteria were intercorrelated for each group. For the junior groups it was found that these ratings intercorrelated highly, suggesting that the juniors did not appear to be able to discriminate between members in regard to these criteria, but rather were influenced in their ratings by the general prestige of the different members of the group. At the senior level, rank correlations between popularity and ability at English, ability to lead in a social situation and ability at English, and ability at organizing and ability at English were not significant at the 0·01 level. As significant correlations were to be expected from among the other criteria, this suggests the development of capacity to discriminate between other children in regard to their possession of these characteristics.

Secondly, it has been shown, in the examination of the essential determinants of leadership in these situations, that from the junior to the senior level there is an increase in the relative importance of the required tools of the situation and of group relationships over prestige as a determinant of leadership. This view gains support from analysis of the reasons given by group members for their choice of persons most suitable as leaders in the two situations. These reasons could be reduced to two categories called 'required tools' (such as 'good at English', 'speaks well' and 'could organize English lessons' for the learning situation and 'knowledge of games', 'able to keep order', and 'could organize a social' for the social situation) and prestige (such as 'popular', 'liked' and 'easy to get on with'). The result of applying χ^2 to test the significance of the differences

between the senior and junior groups on the distribution of the reasons between the two categories, was significant at the 0·01 level. At the senior level the members of the groups stressed reasons related to the required tools of the situation, while the junior groups stressed reasons related to the prestige of members.

It is suggested that the development of leadership differentiation revealed in this study is related to the increasing maturity of children during adolescence as expressed in their developing ability to relate generalizations and concepts to issues in their immediate environment, and it is further suggested that the conclusions reached provide evidence to account for the increased specificity with age of the leaders in the two experimental situations examined.

VII. The effect of revealed differences between the groups on the emergence of leaders

Study of the material obtained from the interview revealed that there were differences between the groups from the different schools which could be summarized as follows:

(*a*) A significant difference in the economic and cultural level of the homes from which both groups from the private school and both groups from the grammar school were drawn.

(*b*) Significant differences in the pattern of group relationships established within the two sets of groups.

(*c*) A significant difference between the private and the grammar school groups in regard to the possession of certain of the required tools for the situations in which the groups functioned.

The differences in regard to the home background of the groups did not result in the leaders at the private school being chosen more from those with the higher home economic status or home cultural level than at the grammar school, nor did the possession by the members of the groups at one school of a greater amount of experience or skill for a particular situation cause differences in the emergence of leaders.

However, there was sufficient evidence to suggest that the differences in the pattern of relationships in the groups at the two schools did affect the pattern of the emergence of leaders in the situations at the two schools. This can be illustrated if we reflect on the interesting fact that, although the leaders in both the senior groups were different in both situations, the correlation between the participation scores in the social and learning situation at the grammar school was 0·18, while the corresponding correlation at the private school was 0·55. Although the senior leaders were different at the private school, the group tended to structure itself in the same way for the two situations. A preference test in this group based on the criterion of 'friends with in leisure time' revealed two cliques of girls (i.e. having no preferences either first, second, or third, for any person outside the clique) and a loose aggregation (i.e. loose connexions with no distinct groupings) among the boys. In the learning situation the girl leader was drawn from one of the cliques of girls which was very active in class work. The other leader was a boy from those of the aggregation whose expressed friendships preferences for girls were directed to members of the clique from which the girl leader was drawn. The other clique of girls, although not actively hostile, made few positive endeavours to help the group towards its goal. In the social situation, the two new leaders, both girls, were again drawn from the active clique of girls, received similar support from the boys as did the previous

leaders and the activity of the other clique was again less than that of other members. As this other clique contained ten members and they occupied the bottom half of the ranking of the participation scores in both situations, the resulting rank correlation reached 0·55, although the leaders were different. On the other hand, at the grammar school, there were no clearly defined cliques in the senior group, merely a number of smaller subgroups of a few members among the girls and an aggregation among the boys. There was no clique pressure to have a member as leader and no support from the boys for any particular group of girls. In this case the group tended to structure itself in an entirely different fashion for both situations. The non-residential—residential difference between the two schools probably accounts for the differences in the pattern of group relationships, and the presence of the difference made it possible to show that a difference in the pattern of group relationships within groups does result in differences in the pattern of the emergence of leaders.

VIII. Summary

A study of the dynamics of the emergence of leaders in two experimental situations was made in natural well-established school groups. The results bear on leadership and group structure as exemplified by adolescent school groups. Examination of what happened when these groups had to handle situations on their own revealed the importance of structure in their sound functioning. The leader-follower dichotomy was shown to apply only to special instances of behaviour in these natural groups, and it was seen that group structure changed with alteration in the urgency of attaining the goal and with variation in the leader's possession of the tool requirements for the situation. The leaders who emerged were, at the senior level, specific to the situation and it was shown that the tool requirements for the situation, group relationships and prestige arising out of other activities were the main determinants in the emergence of leaders. Personality adjustment, age, intelligence, health and home background did not appear to have any bearing on the emergence of leaders in these groups in the situations studied. Different patterns of group relations were found in the grammar and private schools and these affected the emergence of leaders. This difference seemed connected with residence and non-residence rather than with differences in economic background or home cultural level. The differences revealed at the two age levels suggested that the ability to differentiate the requirements for the various leader positions and to identify the persons possessing them develops during adolescence. It is suggested that this development is due to increasing ability, with age, to handle and dissociate concepts. Finally, together with the fact that the developed pattern of group relations played an important part both in the emergence of leaders and in the way in which the group became structured in the situations, the preference tests revealed definite forms of relationship within the four groups. Isolates, near-isolates, stars, cliques, subgroups and aggregates, with their consequent psychological effects, were revealed by analysis of the preferences given by the members of each group. In other words, there are networks of relations between members of groups in addition to teacher-pupil relationships. There is scope for research into the effects of the various positions which members might occupy in this pattern on their ability to make full use of the opportunities of school life, and, where necessary, into the remedies which should be adopted to rectify difficult positions.

(*Manuscript received* 2 *August* 1952)

'EITHER-OR' QUESTIONS IN SERIES[1]

By F. H. GEORGE
University of Bristol

I. *Introduction* (pp. 243–244). II. *Method* (pp. 244–245): (i) *Stimulus materials* (p. 244); (ii) *Subjects* (p. 244); (iii) *Experimental design* (p. 245). III. *Procedure* (p. 245). IV. *The second experiment* (pp. 245–246). V. *Results* (p. 246). VI. *Discussion* (pp. 246–247). VII. *Summary* (p. 247). *References* (p. 247).

I. Introduction

The problems of bias in a series of 'either-or' questions seems to be a matter of considerable interest in the design of experiments, as well as in such particular activities as the compilation of 'either-or' type questionnaires. The assumption that there is an independence between the questions in a series has often, even usually, been made in their statistical treatment. Whitfield (4) has emphasized the fact that biases appear to occur in such questionnaire situations; particularly those involving 'either-or', and in particular, 'yes-no', questions. It should be emphasized that this effect is greatest when the questions involve difficult judgements, or judgements about which the subject feels no certainty.

The problem of such biases, in apparently chance situations, may also have some statistical significance in the interpretation of various experimental, and clinical, situations. The type of guessing behaviour involved has been described by Whitfield as 'randomizing in short sequences', with the frequencies of each response fixed, and *generally* fixed by the subject at approximately equal proportions of each response. The effect on the subject is, in a long sequence, to make responses, on the basis of the whole series being typified by the few immediately preceding responses. Thus a sort of preconceived randomized behaviour will be expected from the subject, at least until he may discover that the sequence is *organized* on some other basis than the random one. This general response may be illustrated by the situation in which a penny is tossed, say three times: the typical response to two successive heads is to believe that the 'chances' are that it will be tails next time. Similar intuitive notions of probability, or the 'random' distribution of events, which generally lead the subject to shun extreme frequencies, appear to be a characteristic of human behaviour.

The experiments, and the one repeated experiment, which this paper describes were designed to see whether this particular sort of biased behaviour could be exhibited under experimental circumstances. Could a series be biased in such a way that an 'absolute' judgement, at a level of comparative uncertainty, be made one way on one occasion, or in one series, and the opposite way in another series.

A brief word of explanation should, perhaps, be given of the immediate background of this experiment. The first experiment was done in 1946–7 (1), and detailed results were

[1] This paper represents the results of an experiment originally suggested by J. W. Whitfield, formerly of the University of Cambridge and now of the University of London. The experiment was carried out by me, originally in the Department of Experimental Psychology, University of Cambridge, in 1946–7, and repeated, by myself, at the University of Bristol in 1952. I wish to express my appreciation to Mr Whitfield, and many others, who have contributed in various ways.

mislaid for some years. The author decided to repeat the experiment, or the weight-lifting part here quoted of the original series of experiments. During the repeated experiment (both experiments were carried out under identical conditions) the original results came to light, and thus both sets of results are given in this paper. The complete agreement between the two experiments, after an interval of some five or six years, is not without interest.

In the broader historical context, these experiments may be seen as closely related to the work of the early psycho-physicists. In particular, the work of Pratt (2) and Woodrow (5) may be regarded as a more general approach to the subject of ordered stimulus-response relations. Their principal concern was with 'time-error' and, of course, this, in particular, involves change of thresholds in relative intervals of uncertainty. Pratt quotes Guilford and Park's experiments as showing that interpolated weights in a series of comparison pairs change the threshold in the direction of the interpolated weight, i.e. higher, if following a very heavy weight and vice versa. Pratt's conclusions ((2), p. 297) are of special interest and relevance.

The preponderance of 'greater' judgements in such comparison-pair situations has also been well established, where this is subject to symmetry of the variables about the standard weight.

Woodrow (5) made his sets of comparisons in terms of both constant and variable standards of comparison. He found, for example, that the difference threshold, when taken as a percentage of the standard, was about 27·5% greater for the varying standard than for the average of two fixed standards. Woodrow believes that the first stimulus in each case is followed by an attitude of expectancy, or set, towards the second.

The present experiments dwell on other aspects of the judgements, and an attempt has been made to extrapolate from the work of the Pratt and Woodrow kind, to say that when a series of paired comparisons enters the uncertainty range, then what is revealed is the subjective estimate of probability distribution in the subject.

II. Method

Using one of the psycho-physical methods; two series, of paired comparisons, of weighted tins were presented to the subjects, who were asked in each instance to say which tin was the heavier. The series were biased, one in favour of an incorrect, and one in favour of a correct, answer at the third comparison pair (the third was in fact the same in each series), and the effect of the bias was estimated statistically.

(i) *Stimulus materials*

Four cylindrical tins, identical in appearance; 5 in. tall, and $2\frac{3}{8}$ in. diameter, were wholly covered in black cartridge paper and weighted inside so that the overall weight of each tin was 180, 195, 210 and 225 g., respectively: they were called tin *A*, *B*, *C* and *D*, respectively.

(ii) *Subjects*

The subjects, sixty in number, were largely undergraduates of the University of Cambridge in the early experiments, and largely undergraduates of the University of Bristol in the repeat of the first experiment. Some non-academic civilians were, however, used in both experiments.

(iii) *Experimental design*

All the subjects were put through a standard procedure of weight lifting, using the standard tins in the form of paired comparisons. There were two series of ten paired comparisons. One series was presented in the order *D–A*, *D–A*, *B–C*, *A–D*, *D–B*, *A–C*, *D–A*, *A–D*, *D–A*, *A–D*; which involved as correct judgements: first, first, second, second, first, second, first, first, second, first. The other series (presented first) was presented in the order *A–D*, *A–D*, *B–C*, *D–A*, *B–D*, *C–A*, *A–D*, *D–A*, *A–D*, *D–A*; with correct judgements: second, second, second, first, second, first, second, first, second, first. It will be seen that all the judgements in both the series are fairly easy (this was ascertained by a pilot experiment), except the third in each series: (*B–C*), which is the most difficult judgement, and is also the same for each series. All the other pairs in each series are the exact opposite of each other, thus giving a standardized background to the critical comparison pairs.

Thus, in the above situation, there is in the two series, an attempt to bias the answer to the third question, by putting it in a series of two different kinds. Having said 'second' 'second', on fairly easy discriminations, and then being faced by a more difficult discrimination, does the subject say 'second' or 'first', when the correct judgement is actually still 'second', but by a much narrower margin than in the first two questions? As a control group the second series had the same third question with the correct answer: 'second', following after two fairly obvious answers of 'first'.

Briefly, the design is intended to answer this question: does a subject, when confronted by an 'either-or' situation in a series (such as in a questionnaire), give absolute judgements, or is there some distortion, effected by the place of the particular judgement in a series of such judgements?

III. Procedure

Each subject was treated individually. The subject was told that he would be asked to make judgements between pairs of weights—no two weights were the same in any pair—and in each case the correct answer was either 'first' or 'second' (i.e. either the 'first' or the 'second' is the *heavier*). They would be asked to test the pair of weights successively with the same hand, under conditions which were as nearly standard as possible.

The tins were concealed by a screen from the subject, who had a large sheet of white cardboard placed on the other side of the screen. There was a circle drawn on the cardboard, and this was kept at a constant distance from the subject who was seated in a chair. The tins were placed by the experimenter, and the subject was asked to lift the weights, and merely say 'first' or 'second', after the second tin had been lifted. The time intervals between presentations were standardized as nearly as possible.

IV. The second experiment

The second experiment, which was done just after the original weight-lifting experiment, used a design that was identical in all operative respects, with the weight-lifting experiments, except in so far as it used light stimuli instead of weight stimuli.

The subjects were asked to view the light stimuli in a completely darkened room, and make verbal judgements as to whether the 'first' or 'second' of pairs of light-stimuli was the longer in length. The sets of pairs were identical with the weight-lifting sets in the

distribution of correct answers. Thus exactly the same bias was introduced into a series of visually determined stimuli.

The results obtained (see Table 2) are comparable with the results obtained in the weight-lifting situation and have virtually the same degree of statistical significance.

The apparatus was a simple box with a circular glass disk $\frac{1}{2}$ in. in radius, behind which was a 5 amp. bulb which was lit when a metal connector passed over a metal sector of a rotating circular strip. The speed of rotation and/or the point where the connector crossed the metal sector could be varied to increase or decrease the interval of time during which the bulb was lit. The time intervals used which were the equivalent of tins *A*, *B*, *C* and *D* were 1, $1\frac{1}{2}$, 2 and $2\frac{1}{2}$ sec., respectively. Thirty subjects were used, and they observed the lighted circular disk from a standard distance of 10 ft. The two sets of ordered pairs were of exactly the same form as in the first experiment.

Table 1. *Correct and incorrect responses to biased comparison pairs (visual discrimination)*

	Correct	Incorrect
Series I. Bias to incorrect response	15	15
Series II. Bias to correct response	29	1

Table 2. *Correct and incorrect responses to biased comparison pairs (weight-lifting)*

	Exp. I		Exp. II		Combined results	
	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect
Series I. Bias to incorrect response	17	13	16	14	33	27
Series II. Bias to correct response	27	3	28	2	55	5

V. Results

The results shown in Tables 1 and 2 appear to be conclusive under the experimental conditions as stated. The results for Exps. I and II, and the combined results, refer *only* to question 3 in both series. The results of all these sets are statistically significant, regardless of the statistical methods employed. The methods which show significance are Whitfield's (3) method for recording probability, or by use of *t*-values, and thus comparing the differences between the two scores of the group, by a suitable scoring method (1 point for a correct reading, nil for an incorrect). It is found that *t*-values are significant at a 5% level of confidence. Inspection of the results, indeed, is probably sufficient in this particular case.

The statistical significance of these results means that the bias towards an incorrect answer was effective in the one series, as compared with the series in which there was a bias towards the correct answer. This is definite statistical evidence for a serial effect in the asking of 'either-or' questions, and suggests the need for the calibration of such questionnaires.

VI. Discussion

The above results are clear in stating, under certain specified conditions, that the answer to a particular 'either-or' question is dependent on its position in the series, as well as upon other factors. In fact the effect of the biased series will presumably increase, as the subject's feeling of doubt increases. At the other extreme, *certainty* of the answer reduces the bias effect to nil.

Thus the experiment corroborates the suggestions made by Whitfield, and shows that, at least at the 'guessing level', serial bias effect occurs. This effect may be taken to illustrate the subject's intuitive sense of probability, which may of course be in conflict with the statistical facts. To be specific, there appears to be an assumption that, in questionnaires involving 'either-or' type questions, the distribution of questions will be in a nearly random manner, and this assumption is persisted in, until a form of 'pattern' (e.g. all 'yes' answers) occurs, when a new assumption is made. It would appear that the assumption of random distribution, at the level of difficult judgements, was sufficient to account for the foregoing results.

How far these results may be generalized is a matter for further discussion, as is the question of individual variation. However, it seems reasonable to assume that whenever ('either-or') dichotomous judgements, or even trichotomous judgements, are made, especially under circumstances which diminish the level of certainty of judgement, then these effects may be expected to occur.

The work of Pratt & Woodrow is almost certainly relevant to these experiments, but it is the opinion of the experimenter that, for judgements in the uncertainty interval, the sense of either kinesthetic or visual discrimination is a variable of the second-order of smallness, while it is the estimate of total distribution of answers that dominates.

One report from a subject in the first experiment is worth recording. At the end of the experiment the subject approached the experimenter in all sincerity, and said that he did not imagine the experiment was very fruitful, since he, the subject, found all the judgements much too easy. In fact, this particular subject, who had, of course, no knowledge of the purpose of the experiment, had made only one mistake out of the twenty judgements: the mistake to which he was experimentally biased.

VII. Summary

One experiment in weight-lifting was carried out twice, with thirty subjects on each occasion and another in visual discrimination was carried out once with thirty subjects. The subjects were presented with a standardized situation, in which they were asked to make judgements as to which tin, of a pair, was the heavier or which light-flash of a pair was the longer: the subject merely said 'first' or 'second' to each of the twenty paired comparisons. The twenty pairs were arranged in two sets of ten, and with the exception of the third judgement, the pairs of the second set were the same as the pairs of the first set, in reversed order. The third judgement in each set was the crucial one: in the one set the subject was biased towards saying 'first', and in the other towards saying 'second', although 'second' was the correct answer in both series. It was found that a highly significant number of subjects made the error to which they were experimentally biased.

References

(1) George, F. H. (1947). Two experiments and their results. Unpublished.
(2) Pratt, C. C. (1933). The time-error in psychological judgments. *Amer. J. Psychol.* xlv, 391–416.
(3) Whitfield, J. W. (1947). Rank correlations between two variables, one of which is ranked, the other dichotomous. *Biometrika*, xxxiv, 292–6.
(4) Whitfield, J. W. (1950). The imaginary questionnaire. *Quart. J. Exp. Psychol.* ii, 76–87.
(5) Woodrow, H. (1933). Weight-discrimination with a varying standard. *Amer. J. Psychol.* xlv, 391–416.

(*Manuscript received* 9 *December* 1952)

PROBIT ANALYSIS OF SUBJECTIVE REACTIONS TO THERMAL STIMULI—A STUDY OF RADIANT PANEL HEATING IN BUILDINGS

By F. A. CHRENKO

Environmental Hygiene Research Unit, Medical Research Council Laboratories, Hampstead

I. *Introduction* (p. 248). II. *Methods* (pp. 249–252): (i) *Radiation calculations* (pp. 249–251); (ii) *Instrumental measurements* (p. 251); (iii) *Subjects and experiments* (p. 251); (iv) *Analysis of results* (p. 252). III. *Results* (pp. 252–255). IV. *Discussion* (pp. 255–256). *References* (p. 256).

I. Introduction

In this paper the term 'radiant panel heating' refers to the use of flat surfaces maintained at relatively low temperatures, that is, not exceeding about 150° F., as the primary source of heat in a room. In modern panel-heating installations, engineers have tended to place the source of heat in the ceiling. This appears to be particularly convenient when multi-storeyed buildings are considered since the heat from the ceiling of one room will warm the floor of the room above. The vogue for this type of ceiling heating is increasing but caution is necessary as, in the interests of economy, there is nowadays a tendency to lower the minimum permissible height of ceilings. It is well known that excessive warming of the head is undesirable and if the source of heat is to be brought nearer to the head the risk of causing discomfort to occupants is increased. British heating engineers have been aware of this risk but they have not been able to assess it, and they sought guidance as to the maximum surface temperature of panels, which, with varying dimensions, embedded in ceilings of different heights, could be used without risk of causing discomfort to occupants of buildings.

An experimental investigation of the subjective reactions of persons exposed to radiation from heated ceilings was recently made in order to obtain the required information.

Since one was dealing primarily with warming of the head it was appropriate to calculate the intensity of the radiation reaching the head from the ceiling with reference to a point at head level. However, the sense of temperature is not apparently uniform over the head. The scalp, for instance, is said to be relatively insensitive to changes in temperature, while other regions appear to be much more sensitive to thermal stimuli. The intensity of radiation emitted from a heated ceiling reaching (*a*) the forehead, and (*b*) the top of the head were considered, together with the mean radiant temperature at a point or small sphere at head level. Each of these three radiation parameters takes into account the dimensions of the heated surface, its temperature and the distance from the head of the subject.

It remained to choose the parameter of radiation which was best for the purpose of predicting the risk of discomfort, and to decide on the limits to be set. Finally, one could then calculate the maximum desirable temperatures of the surfaces of panels of various sizes embedded in ceilings of different heights for the use of the practical heating engineer.

II. Methods

The experiments were made at the Building Research Station, Garston, in a room about 12×18 ft. in plan with a ceiling of adjustable height. The temperature of the walls, floor and ceiling can be controlled separately. The rate of ventilation and the temperature of the air in the room can also be controlled. The heated portion of the ceiling was equivalent to a panel of dimensions 11×16 ft. In addition, an electrically-heated panel, 6×4 ft. in size, was assembled in the room and supported in such a way that it could be easily and quickly raised and lowered.

(i) *Radiation calculations*

The intensity of radiation emitted from a heated ceiling reaching (*a*) the forehead, and (*b*) the top of the head, and also (*c*) the mean radiant temperature at head level, were calculated from the following expressions which were derived by my colleagues, Mr J. McK. Ellison and Dr E. M. Wasilewska.

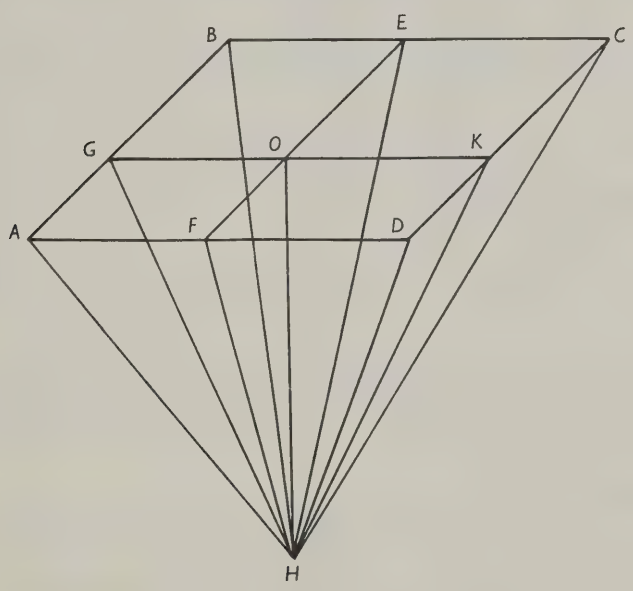

Fig. 1.

Consider Fig. 1. Let $ABCD$ represent a rectangular ceiling panel. Let H represent a point below the panel. From H the perpendicular HO meets the panel at O. The line GOK is parallel to AD and BC, and cuts AB and CD at G and K respectively. The line FOE is parallel to AB and CD, and cuts AD and BC at F and E respectively.

Let $OH = h$, and

$$\left.\begin{aligned} AF = GO = BE = x_1 \\ FD = OK = EC = x_2 \\ AG = OF = DK = y_1 \\ GB = OE = KC = y_2 \end{aligned}\right\} \text{ measured in feet.}$$

Let

$$\left.\begin{array}{l}\angle EHF=\phi_1\\ \angle AHB=\phi_2\\ \angle DHC=\psi\\ \angle BHC=\rho\\ \angle GHO=\theta_1\\ \angle AHD=\theta_2\end{array}\right\} \text{ measured in radians.}$$

Let the distances from the corners of the panel to H be p_1, p_2, p_3 and p_4, i.e. $p_1=BH$; $p_2=CH$; $p_3=DH$; $p_4=AH$.

Let $\sigma=$Stefan-Boltzmann constant $(1{\cdot}72\times10^{-9})$,

$T_1=$surface temperature of panel expressed in degrees Fahrenheit absolute,

$e_1=$emissivity of panel surface,

$dA_2=$an infinitesimal surface area below the panel.

Radiation on forehead. The radiation reaching the forehead of a subject sitting beneath the panel is obtained in the following way.

If dA_2 represents a small receiving surface parallel to and facing a vertical plane containing AB, the effective radiating area of the panel with respect to dA_2 (and the forehead of the subject) is therefore the rectangle $ABEF$.

It can be shown that the radiant energy emitted by the panel on to the vertical section dA_2 is equal to

$$\frac{\sigma e_1 T_1^4 dA_2 h}{2\pi}\left(\frac{\phi_1}{h}-\frac{\phi_2}{\sqrt{(x_1^2+h^2)}}\right), \tag{i}$$

the whole expressed in B.Th.U. per hour.

The expression may also be written as

$$\frac{\sigma e_1 T_1^4}{2\pi}\left(\phi_1-\phi_2\cos\theta_1\right) \quad \text{B.Th.U./hr./sq.ft.}$$

Radiation on top of head. The radiation reaching the top of the head of the subject is obtained by considering dA_2 as a small receiving area parallel to the panel. The effective radiating area with respect to dA_2 is $ABCD$, and it can be shown that the radiant energy emitted from $ABCD$ is equal to

$$\frac{\sigma e_1 T_1^4 dA_2}{2\pi}\left[\frac{x_1}{\sqrt{(x_1^2+h^2)}}\,\phi_2+\frac{x_2}{\sqrt{(x_2^2+h^2)}}\,\psi+\frac{y_1}{\sqrt{(y_1^2+h^2)}}\,\theta_2+\frac{y_2}{\sqrt{(y_2^2+h^2)}}\,\rho\right]. \tag{ii}$$

Mean radiant temperature. The third radiation parameter depends upon the solid angle subtended by the whole panel at a point below it.

It can be shown that the solid angle Ω in steradians at the point H is given by

$$\Omega=\tan^{-1}\frac{x_1y_2}{hp_1}+\tan^{-1}\frac{x_2y_2}{hp_2}+\tan^{-1}\frac{x_2y_1}{hp_3}+\tan^{-1}\frac{x_1y_1}{hp_4}, \tag{iii}$$

and when the point H is below the centre of the panel so that $p_1=p_2=p_3=p_4$:

$$\Omega=4\left(\tan^{-1}\frac{x_1y_1}{hp}\right). \tag{iv}$$

Now suppose that the surface temperatures of the panel and other surfaces of the room are respectively T_p and T_s degrees Fahrenheit absolute. If the solid angle subtended by

the panel at head level is $x\%$ of 4π steradians, the solid angle subtended at a given point at head level by the other surfaces is $(100-x)\%$ of 4π steradians. Let T_m represent the mean radiant temperature at the given point. Then

$$T_m^4 = \frac{x}{100} T_p^4 + \left(\frac{100-x}{100}\right) T_s^4, \qquad \text{(v)}$$

and the mean radiant temperature expressed in degrees Fahrenheit may be obtained from T_m^4.

(ii) *Instrumental measurements*

The mean radiant temperature at head level was also measured directly with a sensitive thermopile.

Thermocouples were used to measure the temperatures of the surfaces of the ceiling, walls and floor of the test room. The humidity of the air in the room was obtained from readings taken with a whirling hygrometer. The atmospheric temperatures were also measured with thermocouples.

All experiments were carried out in calm air, and silvered kata-thermometers were used for estimations of the speed of air movement.

(iii) *Subjects and experiments*

There were two groups of subjects: (1) the writer and four of his colleagues; and (2) 152 members of the staff of the Building Research Station.

Subjects, dressed in normal clothing, were comfortably seated in the experimental room, at first suitably screened from all overhead radiation, and then they were exposed to the radiation from the heated panel or ceiling, at various heights and temperatures, for periods of 30 min. Since all environmental conditions were kept constant during a given experiment, the thermal stimulus applied to the subject was the overhead radiation.

The subjects were questioned concerning their sensations of warmth of (*a*) the head, (*b*) the feet and (*c*) the rest of the body, on the basis of the following scale (1), the assessment of the sensations being made by the experimenter:

Warmth sensation	Numerical mark assigned
Much too warm	+3
Too warm	+2
Comfortably warm	+1
Comfortable	0
Comfortably cool	−1
Too cool	−2
Much too cool	−3

Preliminary experiments had clearly shown that, even though his head remained 'comfortably warm', the subject experienced unpleasant sensations while he was exposed to overhead radiation. The quantitative assessment of discomfort of this nature in terms of an arbitrary scale was a difficult matter, first, because the precise nature of these sensations was not known, and, secondly, because other workers had found that subjects expressed reluctance or inability to commit themselves when asked to assess their discomfort using such a scale. The subjects were therefore asked whether the environment was 'unpleasant' or not.

(iv) *Analysis of results*

The best available method of analysis of such quantal responses, at the time of the experiments, was that of probit analysis, which appears to be based on the earlier work of Fechner (5). Convenient and relatively simple computational procedures for probit analysis[1] are described by Finney (6, 7) and Emmens (4).

In this investigation, the percentage frequencies of occurrence of 'unpleasantness' for given values of the radiative thermal stimulus were converted into probits. These were plotted against the values of the thermal stimulus, and regression lines were calculated in the usual way. Whether there was satisfactory agreement or not between a given regression line and the data upon which it was based was determined from the significance or otherwise of χ^2 as obtained during the cycle of computations for each regression line. This technique was used in the first place to ascertain which of the three radiation parameters was most closely associated with the incidence of discomfort. Secondly, the choice of a suitable radiation parameter and probit regression line enabled a quantitative assessment of the risks of producing discomfort with a known thermal stimulus to be made. Thirdly, an assessment of the risk of producing uncomfortable warmth of the head when the subject was exposed to overhead radiation was obtained. It has already been stated that the warmth of the head of the subject was assessed on the basis of the scale of warmth, and where the recorded vote was +2, i.e., 'too warm', it was clear that the subject's head was uncomfortably warm. For a given value of the radiative stimulus, the number of persons with uncomfortably warm heads was counted and converted into a percentage of the total number exposed, and this percentage was then converted into a probit. In other words, the quantitative scale of warmth was used in order to obtain information of a quantal nature. This information could have been obtained by direct questioning, but the former method was preferred as further analysis of subjective reactions then became possible. Details of this are given elsewhere (3).

III. Results

The mean *overall* warmth vote for all the experiments was nearly equal to zero on the arbitrary scale of warmth, i.e. 'comfortable', and over 90 % of all these votes were within the comfort zone, that is, they ranged from 'comfortably cool' to 'comfortably warm', or between −1 and +1 on the scale of warmth. Localized sensations of warmth of the head and feet varied considerably in these experiments, but the range of sensations, whether overall or localized, was not sufficient to permit statistical analysis using the product-moment correlation method. The word comfort is here used in a restricted 'thermal' sense. Discomfort, on the other hand, is used to denote unpleasant sensations experienced by the subjects. These sensations included thermal and other sensations, for instance those associated with gradients of thermal sensation.

Space does not permit a detailed presentation of the results of all the experiments which

[1] If, when a stimulus is applied to the subject, the response is quantal (i.e. all-or-nothing), and if the proportion (P) of positive responses is a function of the stimulus (x) then, in probit analysis (the word probit is derived from 'probability unit') it is assumed that:

$$P = \frac{1}{\sqrt{(2\pi)}\,\lambda} \int_{-\infty}^{x} e - \frac{(x-\mu)^2}{2\lambda^2}\, dx,$$

where λ = standard deviation of the distribution of the thresholds; μ = mean value of the thresholds. Tables are available from which percentages may be converted into probits.

were made, and it must therefore suffice to state that preliminary plotting of the results indicated that, of the three radiation parameters investigated, one was more closely associated with the risk of discomfort than the other two, and this was confirmed by subsequent probit analysis.

This radiation parameter, based on equations (iii), (iv) and (v), was the elevation of the mean radiant temperature at head level (° F.) due to the heated panel or ceiling, and it was obtained in the following way. First, the mean radiant temperature at head level in that part of the room where the subject was seated and screened from all overhead radiation was calculated. Secondly, the mean radiant temperature at head level where the subject was exposed to overhead radiation was calculated, and from the figure thus obtained the first mean radiant temperature was subtracted. The resulting difference is a measure of the radiative stimulus applied to the head of the subject, since all other environmental factors were constant during a given experiment.

In other words, the percentage frequency of occurrence of unpleasant sensations was, for all subjects, more closely associated with the radiation received by the whole of the head than with the intensity of radiation received either by the forehead or the top of the head.

This was investigated in another way. Another series of experiments was carried out with the subjects of group 1. Instead of sitting beneath the centre of the panel (as in all previous experiments in which that was used), the subjects sat at a distance of 5 ft. from the centre of the panel. The results gave no indication that the incidence of unpleasant sensations varied with the orientation of the face of the subject with respect to the heated surface overhead, although with the same elevation of mean radiant temperature there was more intense radiation on the forehead than when the subject sat under the centre of the panel. Moreover, scrutiny of the remarks made by subjects who had voted that a given environment was unpleasant showed that symptoms were referred to the forehead in some instances, to the side of the face in others, or even to the back of the neck. A bald subject in group 2 did not, as might have been expected, refer his unpleasant symptoms to the top of his head, but instead referred them to his forehead.

The results might have been different if high-temperature sources of radiation had been used in the experiments, but it was neither necessary nor desirable to produce intolerable conditions.

It should also be noted that, although some of the subjects experienced uncomfortable warming of the head, the term 'too warm' does not denote gross discomfort. The localized heating of the head was not then severe.

The results of probit analysis of one series of experiments with the heated panel with the subjects of group 1, and another series of experiments with the heated ceiling on the subjects of group 2, are given in Table 1. Considering the 152 subjects of group 2, the results show that the median value of the radiative stimulus causing unpleasant sensations to be experienced by the subjects was 8·3° F. If, therefore, a reasonably large group of subjects is exposed to overhead radiation so that the elevation of mean radiant temperature due to the heated ceiling is 8·3° F., there is a risk that 50% of them will experience unpleasant sensations. The upper and lower fiducial limits of the response under these conditions have been calculated, and they are respectively 57 and 43%.

The probit regression equation for group 2 is illustrated in Fig. 2. The probit of the response of the subjects, i.e. the unpleasant sensations experienced by the subjects on

exposure to radiation from a heated ceiling, is shown on the ordinate, and the radiative stimulus, i.e. the elevation of mean radiant temperature at head level (° F.) due to the heated ceiling, is on the abscissa. The upper and lower curves of the fiducial limits of the response have been calculated and are drawn in the figure.

Table 1. *The relation between the probit of the frequency of occurrence of unpleasant sensations and the elevation in the mean radiant temperature (° F.) at head level due to the radiant heat from the ceiling*

(b and a are the constants in the probit regression equation, $Y = bx + a$, where Y is the probit of the response and x is the elevation of mean radiant temperature in ° F.)

	Subject group 1. Experiments with heated panel	Subject group 2. Experiments with heated ceiling
b	0·182±0·039	0·186±0·037
a	3·74	3·47
Median value of stimulus (° F.)	6·93±1·19	8·26±0·96
χ^2	8·96	2·71
Degrees of freedom	6	6

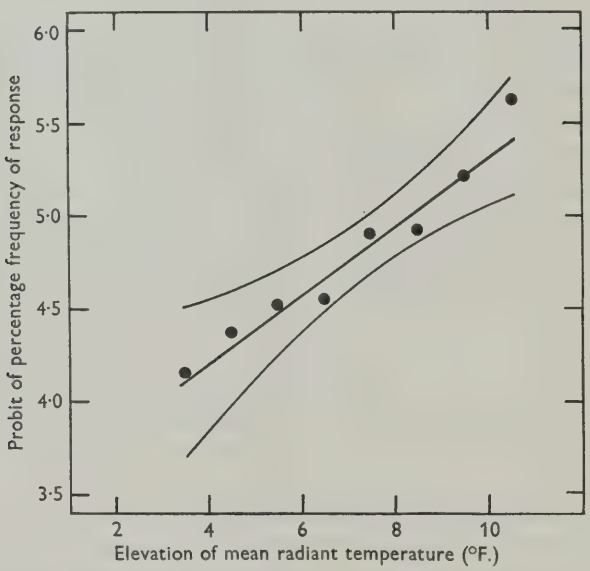

Fig. 2. The probit of the response of the subjects of group 2, plotted against the elevation of mean radiant temperature at head level due to the heated ceiling showing the probit regression line and the 5% fiducial band.

The risk of producing uncomfortable warmth on the head of the subjects of both groups was obtained from another probit regression equation:

$$y = 0{\cdot}13x + 3{\cdot}25,$$

where y is the probit of the percentage of occasions in which the head of the subject was uncomfortably warm (i.e. +2 or more on the arbitrary scale of warmth) and x is the elevation of mean radiant temperature (° F.) at head level, due to the radiation overhead.

Comparison of this equation with the data given in Table 1 shows clearly that, for a given increase in mean radiant temperature at head level, the chances of evoking

unpleasantness are greater than those of over-heating the heads of the subjects. For instance, an increase in mean radiant temperature of 8·3° F., which would give a 50% incidence of unpleasant sensations in group 2, would give a 25% incidence of over-warming of the head.

This suggests that there are other causes of unpleasant sensations apart from over-heating of the head. For example, the head of a subject may be comfortably warm while his feet may be comfortably cool, yet the combined effect may be, and often was, unpleasant. Such gradients of sensation were found to be important causes of unpleasant reactions.

IV. Discussion

There are certain practical advantages to be derived from the finding that the elevation of the mean radiant temperature at head level due to the radiation overhead was the best parameter of radiation of those investigated. In the first place, it is comparatively simple to calculate the mean radiant temperature at any point in a room if the temperatures of the internal boundary surfaces of a room are known. Secondly, if the surface temperatures of the heated panel or ceiling and other surfaces in the room are not known, it is easy to measure the mean radiant temperature directly by means of a radiometer, or to obtain a good estimate of the radiant temperature with the easily constructed blackened-globe thermometer described by Bedford (1). Thirdly, heating engineers are (or should be) familiar with this particular radiation parameter and the methods by which it is measured.

It has been asked whether the designer of a heating installation should plan for the so-called 'average man', or for the trained subject, and the assumption is made that the sensitivity of the former is low while that of the latter is high. The answer is that one should plan for neither one nor the other. For the designer of a panel-heating system, the sensitivity of the 'average man' is represented by the median value of the threshold (value of radiation evoking discomfort), and since the median threshold stimulus corresponds to a 50% risk of producing discomfort it cannot be used as a criterion for design purposes.

Bedford (1) took as the comfort zone that range of conditions in which not less than 70% of the subjects were 'comfortable'. In that range 86% of the responses ranged from 'comfortably cool' to 'comfortably warm', so that there was a risk that 14% of the subjects would be uncomfortable.

It does not follow that environments which, under experimental conditions in which subjects were asked to express how they felt, caused 14% of the subjects to state that they found the environment unpleasant, would necessarily provoke unsolicited complaints under other circumstances. It is likely, however, that environmental conditions which caused 50% or more of our subjects to feel unpleasant would provoke strong complaints from many people other than experimental subjects, and observations made in the field recently tend to support this view. For instance, in a part of one factory heated by radiant panels, where such conditions had prevailed, the operatives took the simple but effective step of shutting off the supply of heat to the offending panels.

For practical purposes it was decided arbitrarily that the maximum desirable elevation in the mean radiant temperature is 4° F., i.e. that the risk of discomfort should be about 20% and the risk of overwarming the head about 10%. Tables of maximum desirable temperatures of heated ceilings and panels at different heights have been calculated on

this basis (3). For example, if a room is warmed by a single rectangular ceiling panel of dimensions 12 by 8 ft., the maximum desirable surface temperature of the panel is 103° F. if the ceiling is 12 ft. above the floor, 91° F. if the ceiling height is 10 ft., and only 82° F. if the height of the ceiling is reduced to 8 ft.

This investigation was initially possible only because Dr F. M. Lea, Director of Building Research, generously allowed the unique Controlled-Weather House at the Building Research Station, at Garston, to be used for our experiments, and granted permission for his staff to act as subjects. The success of the experiments was due, in no small measure, to the active assistance and co-operation of numerous members of the staff, and to them, and especially to those who kindly volunteered to act as subjects, the author and his colleagues wish to express their sincere thanks.

Dr T. Bedford, Director of the Environmental Hygiene Research Unit, gave useful advice and constructive criticism throughout the investigation, and the author's sincere thanks are due to him, and to his colleagues Miss B. E. Tredre and Messrs R. C. Houslip, G. W. Spicer and D. Turner, who all patiently acted as experimenters and subjects for many months.

References

(1) Bedford, T. (1936). The warmth factor in comfort at work. *Rep. Industr. Hlth Res. Bd, Lond.*, no. 76.

(2) Bedford, T. (1946). Environmental warmth and its measurement. *M.R.C. (War) Memo.* no. 17.

(3) Chrenko, F. A. (1953). Heated ceilings and comfort. *J. Instn Heat. Vent. Engrs, Lond.*, xx, no. 209, 375–96.

(4) Emmens, C. W. (1948). *Principles of Biological Assay.* London: Chapman and Hall Ltd.

(5) Fechner, G. T. (1860). *Elemente der Psychophysik.* Leipzig: Breitkopf und Härtel.

(6) Finney, D. J. (1944). The application of probit analysis to the results of mental tests. *Psychometrika.* ix, no. 1.

(7) Finney, D. J. (1947). *Probit Analysis.* Cambridge University Press.

(*Manuscript received* 13 *January* 1953)

HUNTER'S TEST OF THE ABSOLUTE AND RELATIVE THEORIES OF TRANSPOSITION

By G. ROBERT GRICE

Department of Psychology, University of Illinois

In an attempt to test opposing implications of the 'absolute' and 'relational' theories of transposition behaviour, Hunter (5) has recently reported a series of three visual discrimination learning experiments with children. The findings were interpreted as supporting the 'relational' interpretation as opposed to the 'absolute' theory which explains transposition on the basis of the development of habit strength and inhibition to the specific training stimuli, and the generalization of these tendencies to other stimuli in the dimension. Two of the studies involved children ranging in age from 59 to 71 months, and thus are somewhat questionable for this purpose because of the possible utilization of verbal cues in solving the problems. The third experiment, however, involved children from 14 to 32 months of age—a range which should be relatively safe for this purpose (1, 8). The purpose of the present discussion is to question the prediction made by Hunter on the basis of the absolute theory. Unless this prediction is correct and necessary, a crucial test of the theory is not provided. In the writer's opinion, the experimental situation considered may well be one in which the opposing theories lead to identical predictions.

In Hunter's design, *A*, *B*, *C*, *D*, *E*, *F*, *G* and *H* represented a series of stimulus objects varying in size such that they fell along an increasing, equal-unit scale of log area. The plan of the experiment was to train on three problems which were then followed by a test. The problems were $D(+)$ *v.* $A(-)$, $F(+)$ *v.* $C(-)$, and $C(+)$ alone. (In one of the experiments, the third problem was $C(+)$ *v.* a neutral stimulus not on the dimension.) The critical test was the presentation of *E v. H*. *D v. A*, *F v. C* and *E v. H*, each maintained the same ratio of areas. The relational theory was interpreted as predicting the choice of *H*, while the absolute theory was interpreted as predicting the choice of *E*. Actually, *H* was the stimulus predominantly chosen.

In deducing from the absolute theory that *E* should be chosen over *H*, Hunter reasoned that *H* should be affected less by generalization from all three of the positive stimuli, *D*, *F* and *C*, since it is further out on the size dimension. This reasoning is correct, as far as it goes, but it fails to consider properly the role of inhibition generalizing from the negative stimuli, *A* and *C*. Spence's version of the absolute theory (9, 10) makes very explicit that it is the *net* excitatory strength or the difference between generalized excitatory or habit strength and inhibition which determines the strength of the tendency to respond to a particular stimulus. It is, of course, this property of the theory which makes possible the prediction of transposition. Hunter recognized this problem, in part, when he contended that the training to *C* alone would 'ensure an excitatory strength at *C* and enable us to disregard inhibitory generalizations from either *A* or *C*'. The latter part of this statement is incorrect, however, as may be seen from the following discussion.

In order to apply the reasoning of Spence's theory to this situation, it has been assumed

that habit strength and inhibition generalize according to the functions employed by Spence (10). Specifically the functions are

$$H = 100 \times 10^{-0.047X^{2.5}} \quad \text{and} \quad I = 60 \times 10^{-0.094X^{2.5}},$$

in which X equals the difference in logarithmic units between the original training stimulus and other stimuli on the dimension. It is further assumed, for purposes of the present illustration, that the steps between the stimuli in the training pairs and in the test pair are represented by one unit on the X scale. A further point which must be kept in mind is that the curve for the growth of habit strength is negatively accelerated and approaches an asymptotic value represented by a value of 100 in this instance. Presumably, this value is actually approached in the usual discrimination learning situation unless the problem is an exceptionally easy one. The generalization gradients produced by training on the two problems *A v. D* and *C v. F* are indicated in Fig. 1. The *net* excitatory strengths (Hull's (4)

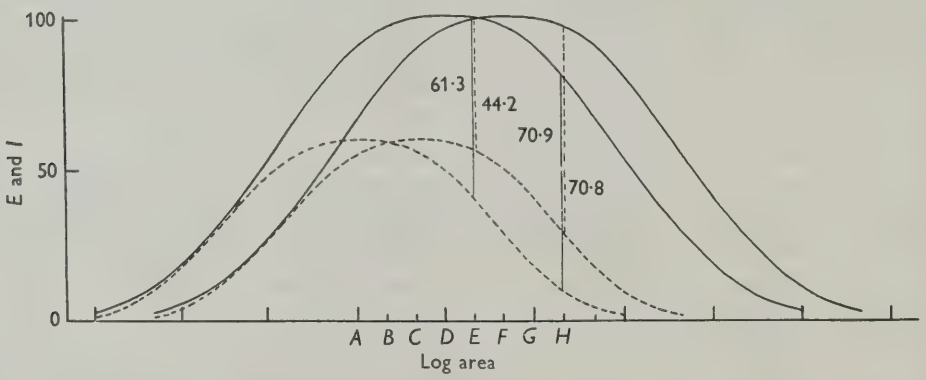

Fig. 1. Generalization gradients produced by training on problems I and II. The numbers to the left of the vertical lines represent the net effective reaction potentials to E and H resulting from problem I, and those to the right represent those resulting from problem II.

'effective reaction potential', ${}_s\bar{E}_r$) accruing to stimuli E and H are indicated by the solid vertical lines for problem I and by the dotted lines for problem II. It will be noted that in each instance the net excitatory strength is greater for H than for E.

Strictly speaking, the theory demands a further step in combining the gradients. Thus, the curves have been summated according to the expressions proposed by Spence (10) and Hull (4).[1] The resulting combined gradients are presented in Fig. 2. The net excitatory strength at H is 55·3 and at E it is 25·1. Therefore, with these assumed gradients, the theory predicts that transposition will occur, and that H will be chosen in preference to E.

But what of the training to stimulus C alone or in combination with a neutral stimulus? Hunter's point that this training will ensure an excitatory strength at C and enable us to ignore inhibitory generalization does not follow from the theory. In fact, it may be seen in Fig. 2, that under the present assumptions habit strength is maximal at C and additional training would not further increase it. This is not essential to the argument, however, for raising the excitatory gradient would still leave H greater than E. There is also a *net*

[1] $E_1 + E_2 = E_1 + E_2 - \frac{E_1 E_2}{100}$, and similarly, $I_1 + I_2 = I_1 + I_2 - \frac{I_1 I_2}{I_{\max}}$.

excitatory strength greater than zero at C as there is to all other stimuli within the range of generalization. The discrimination problem involving the neutral stimulus would consist merely of the extinction of initial response tendencies to the neutral stimulus. The theory as stated by Spence does not involve an assumption that positive training to any stimulus eliminates the gradients of inhibition. In fact, it is essential to the theory that both habit strength and inhibition may be associated with the same stimuli. The combined gradients indicated in Fig. 2 should be approximately the same with or without the additional training to C. In other words, with the generalization functions assumed here, the absolute theory would predict the preference of stimulus H over E just as would the relational theory.

The above deduction is not intended to convey the impression that the absolute theory in its present state could actually predict in advance the specific result of Hunter's experiment. It is intended merely to demonstrate that the theory can, in principle,

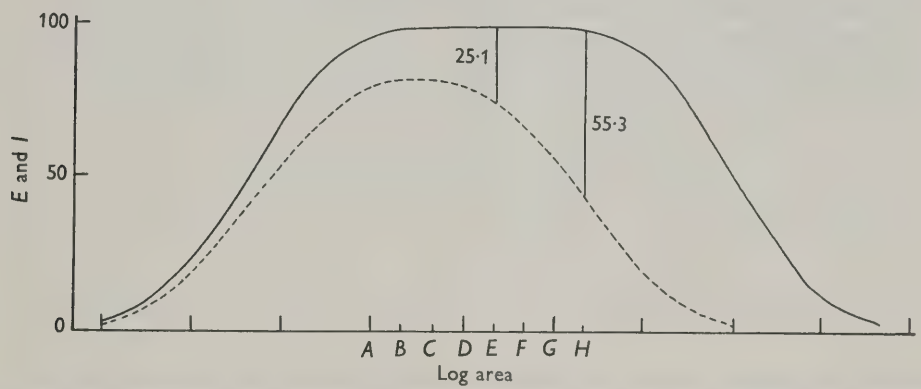

Fig. 2. Combined generalization gradients of excitatory strength and inhibition resulting from training on problems I and II. The numbers by the verticals at E and H represent the theoretical net effective reaction potentials at these points resulting from the combined training.

account for such a result, and that such an experiment does not constitute a test of the opposing theories. Specific predictions of this kind will be possible only when the generalization functions can be based upon adequate empirical data. This is merely one example of the present great need for quantitative laws in psychology. A modest beginning has been made on the task of obtaining the requisite dimensional data (2,3,6,7), but it will probably be some time before adequate quantitative postulates can be stated which will permit such precise predictions. In the meantime, the theory does allow certain predictions of a more general character. For example, in the above situation, if test pairs are taken at various steps progressively further along the dimension, preference for the larger should continue for a time, but a point should be reached at which there would be no preference for either stimulus. Beyond this range, preference for the *smaller* stimulus should be demonstrated. Preliminary confirmation of this tendency has recently been obtained by Ehrenfreund (2). It is of considerable interest that the relational theory has, as yet, no basis for predicting this failure of transposition, and certainly not for predicting its reversal.

References

(1) Alberts, E. & Ehrenfreund, D. (1951). Transposition in children as a function of age. *J. Exp. Psychol.* xli, 30–8.
(2) Ehrenfreund, D. (1952). A study of the transposition gradient. *J. Exp. Psychol.* xliii, 81–7.
(3) Grice, G. R. & Saltz, E. (1950). The generalization of an instrumental response to stimuli varying in the size dimension. *J. Exp. Psychol.* xl, 702–8.
(4) Hull, C. L. (1939). The problem of stimulus equivalence in behavior theory. *Psychol. Rev.* xlvi, 9–30.
(5) Hunter, I. M. L. (1952). An experimental investigation of the absolute and relative theories of transposition behaviour in children. *Brit. J. Psychol.* xliii, 113–28.
(6) Kendler, T. S. (1950). An experimental investigation of transposition as a function of the difference between training and test stimuli. *J. Exp. Psychol.* xl, 552–62.
(7) Kling, J. W. (1952). Generalization of extinction of an instrumental response to stimuli varying in the size dimension. *J. Exp. Psychol.* xliv, 339–46.
(8) Kuenne, M. K. (1946). Experimental investigation of the relation of language to transposition behavior in young children. *J. Exp. Psychol.* xxxvi, 471–90.
(9) Spence, K. W. (1937). The differential response in animals to stimuli varying within a single dimension. *Psychol. Rev.* xliv, 430–44.
(10) Spence, K. W. (1942). The basis of solution by chimpanzees of the intermediate size problem. *J. Exp. Psychol.* xxxi, 257–71.

(*Manuscript received* 31 *March* 1953)

REPLY TO PROFESSOR GRICE

By IAN M. L. HUNTER

Institute of Experimental Psychology, University of Oxford

Of the three visual discrimination studies reported in this *Journal* the first was an attempt to test the relative-absolute issue with the use, as subjects, of children aged 59–63 months. As proposed by Spence, the absolute theory applies only to animals and pre-linguistic children and not to subjects in whom transposition might be effected by the mediation of 'verbal cues' and in whom such behaviour presumably accords with the relative theory. For this reason, Grice (2) seems to suggest that examination of 5-year-old children is irrelevant to the absolute theory. But, surely, the validity of Spence's restriction requires scrutiny as much as any other of his formulations. And this validity can only be tested if it is known whether or not the absolute theory applies to linguistic subjects. Thus, far from being irrelevant to Spence's theory, the study of linguistic subjects is essential for its complete verification. My second study represented an exploratory investigation of stimulus conditions affecting transposition and is not, at present, under discussion. The third study was, in effect, a repetition of the first with pre-linguistic children now acting as subjects. The primary object of Grice's paper is to demonstrate that the absolute theory can, in principle, account for the results of this experiment and that, in design, the experiment itself does not constitute a test of the relative-absolute issue. As far as the writer can see, Grice's reasoning is sound and his conclusion valid.

Although it has now been shown that this test is not crucial, the experiment produced performances by two subjects which seem inexplicable in terms of the absolute theory. These two children, one aged 16 and the other 26 months, made not a single wrong choice throughout training. Thus, while there was excitatory strength at stimuli *C*, *D* and *F*, there was no inhibitory strength at *A* and *C*. So the subsequent preference of *H* to *E* could not be accounted for. Incidentally, the choice of *H* could not be due to an independent size-preference since one subject belonged to group SM and the other to group LA ((3), p. 123). However, there being only two such performances, it would be unwise to over-stress their implications.

The writer agrees with Grice that, as it now stands, the absolute theory has little predictive capacity. It does, however, make one prediction worth mentioning. Suppose that a subject, trained on discrimination *A* (−) *v.* *B* (+), transposes to the pair *B v. C* by acting positively to *C* and negatively to (the formerly positive) *B*. According to the absolute theory, this results from a stronger tendency to act positively to isolated *C* than to isolated *B*. Moreover, the difference in strength between the two tendencies is independent of the simultaneous presence of *C* and *B*. Under favourable circumstances, therefore, any subsequent discrimination which involves acting positively to *C* (or negatively to *B*) will be more easily learned than one involving positive action to *B* (or negative action to *C*). For the relative theory, the preference for *C* over *B* depends entirely on the relation between them and, if this is destroyed by presenting the stimuli in isolation instead of together, the tendency to act positively to *C* will not exceed that

to act so to *B*. The writer(4) has used those divergent predictions as a test of the relative-absolute issue. Since the design, by its very simplicity, is not easily adapted to suit children, rats were used as subjects. The results favoured the absolute theory.

Grice refers to the absolute theory's prediction that, as the similarity between the training and test pair decreases, the frequency of transposition will fall to 50% and then actually fall below it before rising again to level off at chance. Ehrenfreund's(1) 'preliminary confirmation' of this prediction employed, as stimuli, eight different brightnesses, numbered in order from 1 to 8. Training was given on discrimination 1 (−) *v.* 4 (+). The rats transposed to test pairs 2 *v.* 5, and 3 *v.* 6; they showed significantly less than 50% transposition to pair 4 *v.* 7; and they showed about 50% transposition to pair 5 *v.* 8. The reversal, therefore, involved positive action to that same stimulus (i.e. 4) which had been positive in training. This does not, in consequence, unambiguously verify the absolute theory's original prediction. Such a verification could only be produced if the various test pairs did not, as in Ehrenfreund's(1) study, overlap but were separated by discrete steps as, say, in the series 1 *v.* 2, 2 *v.* 3, 3 *v.* 4, etc.

Grice's argument that the relative theory has no basis for predicting failure of transposition would hold only if that theory maintained that the stimulus characteristics effective in controlling behaviour were abstract relations and never absolute characteristics of the training pair. But the subject might well act to both types of characteristic together, i.e. to what might be called stimuli-in-relation. Formally stated, this interpretation of the relative theory would be: when trained to choose between two stimuli of, say different size, the subject acts to the larger or smaller stimulus depending on the direction of training and acts likewise to other pairs of size stimuli with a frequency which is proportional to the similarity (in physical or action-defined terms) of these test pairs to the training pair. In this formulation, which can account for failure to transpose, is combined the relational-action hypothesis of the relative theory and the law of generalization-by-similarity which has been so stressed by the absolute theory and which has proved its worth in many a different type of transfer study.

References

(1) Ehrenfreund, D. (1952). A study of the transposition gradient. *J. Exp. Psychol.* xliii, 81–7.
(2) Grice, G. R. (1953). Hunter's test of the absolute and relative theories of transposition. *Brit. J. Psychol.* xliv, 257–60.
(3) Hunter, I. M. L. (1952). An experimental investigation of the absolute and relative theories of transposition behaviour in children. *Brit. J. Psychol.* xliii, 113–28.
(4) Hunter, I. M. L. (1954). The absolute and relative theories of transposition behaviour in rats. *J. Comp. Physiol. Psychol.* (in the Press).

(*Manuscript received* 14 *May* 1953)

PUBLICATIONS RECENTLY RECEIVED

The Neurophysiological Basis of Mind. By J. C. Eccles. Oxford: Clarendon Press; London: Geoffrey Cumberlege. 1953. Pp. xii+314. 25*s.* net.

This book is based on eight lectures given by its author in Oxford last year. It provides a valuable and fascinating account of much contemporary work in the field of electro-physiology and as such will be welcomed by advanced students of all ages. Particularly important are his own findings with a micro-electrode inserted into the soma of a motor neurone in the cat's spinal cord. The differences of potential between the inside and the outside of such a cell when it is at rest, when impulses are 'back-fired' into it from its own axone, and when it is reflexly excited or inhibited by stimulation of appropriate afferent pathways are analysed with great skill, and their successful recording is one of the author's most recent technical triumphs. He gives evidence not only of depolarization of the surface membrane during excitatory bombardment but also of hyperpolarization when inhibitory synapses are activated. The bearing of this on theories of the nature of central excitatory and inhibitory states is clearly brought out. Some of these theories must now be abandoned and the setting free at synapses of chemical transmitters seems to offer the most plausible explanation.

The title is a little misleading about the book's scope. The first two chapters are wholly concerned with the properties of nerve fibres and muscle fibres. The next four deal with events at motor end-plates, in sympathetic ganglia, and in the grey matter of the spinal cord. The author's own extensive work in this field enables him to write with experience and authority. These chapters will be read with interest by all neurophysiologists. Only in the seventh chapter are we introduced to the immensely more complex happenings in the brain as studied by electrical recording from the cerebral cortex. This is followed by a final chapter in which 'fragmentary and tentative' hypotheses are put forward as possible explanations of how mind could influence brain in willed movements and how brain might act on mind in perception. These are admittedly pure speculation, of necessity unsupported by experiment. The only references elsewhere to mental activities are to the neurological basis of learning and memory. Experiments on the spinal cord lead to the suggestion in chapter VI that plastic changes of long duration (particularly swelling) take place at synapses, whereby their excitatory effect is increased by usage and depressed by disuse.

It is inevitable that a book on so intricate and technical a subject as electro-physiology should make demands on the reader. Publisher and author are to be congratulated on the trouble they have evidently taken to lighten his burden. It might perhaps have been lightened even further by the avoidance of such unusual grammatical constructions as the repeated (but not invariable) use of 'propagate' as an intransitive verb and of 'data' as a singular noun. R.S.C.

Design for a Brain. By W. Ross Ashby. London: Chapman and Hall. 1952. Pp. ix+260. 36*s.*

This book expounds systematically an approach to biological problems using linguistic and mathematical tools similar to those now employed in communications engineering and the construction of servo-mechanisms. The author treats as the central problem the fact that an animal's behaviour may change its pattern from one predictable form to another, and will usually settle into a form which is homeostatic: that is, in which certain essential variables stay within a particular range of values. Starting from first principles, he states this problem in his own terms; and accounts for it by a principle of 'ultrastability'. This asserts that observed patterns will change sharply but randomly when critical states are reached, and consequently will go on changing until a pattern appears in which critical states do not occur. Provided the critical states are related to the essential variables already mentioned, such a system will give adaptive behaviour. Dr Ashby's homeostat, a machine which embodies this principle, is described, and the remainder of the book considers the properties of systems containing numerous 'ultrastable' mechanisms in combination. A mathematical appendix restates some of the arguments formally, the main text requiring no technical knowledge: cross-reference within and between both sections is easy.

Such a summary does little justice to the main virtue of the book, which is the clear and striking development of the terms and the generally stimulating quality of the presentation: the advantages of a precise language cannot be expressed in a review using everyday terms. The main weakness may be typified by Dr Ashby's claim 'to show that ultrastability is not merely plausible but necessary: the

organism *must* be "ultrastable".' The only supporting formal argument is that sharp changes in a sub-system imply the presence elsewhere in the main system of a variable changing instantaneously between constant values: and even this argument contains weaknesses. One is noted by the author, and another is the fact (correctly stated elsewhere but ignored in this connexion) that one variable may be independent of another over one range of values but not over another. It is curious to note that no reference is made to Spence's explanation of discontinuous behaviour without postulating step-functions.

The further assumption of 'ultrastability', that the critical states for changes of field of a sub-system are states of that sub-system itself and not of independent variables elsewhere in the main system, is not justified in any way: and without this assumption neither system will tend to a stable form. Existing theories of learning tend, if anything, to adopt the opposite assumption that changes in behaviour occur at random time but with determinate transition probabilities, and such theories are at least as adequate as this one. In addition, various arbitrary assumptions are involved in applying the abstract theory to the particular case of organisms.

It remains true, then, that new knowledge cannot be obtained simply by adopting a new language. But it would be unfortunate if this overstatement, or dissatisfaction with the author's particular theory, were to obscure the value of the book. Verbal issues still occupy too much of psychology, and the subject would benefit if experimenters were to study this book as an example of the use of language. As an introduction to precise description it is not likely to be bettered. D. E. BROADBENT

The Sensory Order. An Inquiry into the Foundations of Theoretical Psychology. By F. A. Hayek. With an introduction by Heinrich Kluver. London: Routledge and Kegan Paul. 1952. Pp. xxii + 209. 18*s*.

In this book Prof. Hayek, the distinguished economist, puts forward a general theory of mind which forms a contribution to psychology and philosophy of unusual interest. Its similarity to the recent work of Hebb makes it particularly significant at the present moment. Only a brief account can be given here. The author begins by considering the problem of the relation of mind and matter in connexion with the perception of the external world. He holds that the distinction between the phenomenal and physical worlds is fundamentally that the classification of events which is given in sensory experience ('the sensory order') has no exact counterpart in the classification of events adopted by the physical sciences. The events which to our senses are alike or different will not necessarily be regarded by the physicist as alike or different. The task of psychology is therefore, taking the physical world for granted as a starting-point, to reconstruct the process by which the organism achieves the phenomenal classification. This is the work of the nervous system. It is to be regarded as an instrument for classifying the impulses which pass through it. We must postulate a neural apparatus which will differentiate between the impulses passing through it in exactly the same manner in which our sensory experience differentiates between the stimuli. This neural differentiation is ultimately a question of *which* nerve fibres are excited, and the neural apparatus is built up, in the development of the species or the individual, by the establishment of connexions between nerve fibres in which impulses occur at the same time. It is then claimed that the structure of such a neural system of classification will be topologically equivalent to the structure of the system of classification in sensory experience, or rather, that a particular physiological event becomes a mental event in so far as it is part of such a structure, and that what we call 'mind' is the particular order of events in such a structure. In the second half of the book the thesis is applied with great ingenuity to a wide range of important psychological and philosophical problems.

From the psychological point of view Prof. Hayek's theory seems to rest fundamentally on two principles: Muller's law of specific nerve energies and Mill's law of association by contiguity. The first principle, which Helmholtz considered as important for psychology as the principle of gravitation for physics, is applied far more extensively than it was by Muller. This is one of the most provocative and penetrating parts of the book. His use of the second principle is, on the other hand, surprisingly uncritical. A modern statement of it surely requires much more careful handling than Prof. Hayek gives it. The chief weakness of the book, however, lies in the author's failure to apply his theory in concrete detail to empirical findings. He does not attempt, for instance, to account for colour vision or to reconcile his views with the currently accepted theory of hearing. Indeed, a section headed 'Observed Facts for which the Theory Accounts' covers less than three pages, and could only by courtesy be said to contain the record of a single observed fact. He never thoroughly examines 'the sensory order', and even leaves one in doubt as to what he means by it. Despite these defects, the book possesses a range and subtlety which should be widely appreciated. R. B. JOYNSON

Un Nouveau Phénomène Psychosomatique. By Caleb Gattegno and Alphonse Gay. Neuchâtel: Delachaux and Niestlé. 1952. Pp. 82. No price given.

At the beginning of the century Dr Vittoz of Lausanne discovered that a slight rhythmical throbbing or pulsation was often to be observed on the surface of the forehead. It seemed to be unconnected with any known physiological reaction and might, he suggested, be of value in the diagnosis of mental disorders. Except for some work by d'Espiney in the twenties the phenomenon appears to have escaped further notice until recent years, when it has been investigated by the authors of the present brief report. They state that it is not restricted to the forehead and suggest that its immediate cause lies in minute changes in the elastic tension of the capillaries, ultimately dependent on the action of the sympathetic nervous system. If this is so, it is clearly to be placed in the same general category as the PGR and a comparison of the two, which the authors do not make, would be useful.

In their investigations they employ a sensitive electronic apparatus, named the gayograph after its inventor. They print numerous examples of the rhythms which this instrument records and claim that various psychological states, normal no less than abnormal, and intellectual as well as emotional, display characteristic rhythms. They consider that this is a promising field of inquiry for both theoretical and applied psychology. The small number of subjects who appear so far to have been examined, and the general trend of research on peripheral indications of central states, make one cautious in estimating the ultimate significance of this work. But whether or not 'the Vittoz phenomenon' proves to be more than a curiosity, the authors are to be congratulated on their enterprise and enthusiasm. We shall await news of their further researches with considerable interest.

R. B. JOYNSON

The Idea of Progress: A Revaluation. By Morris Ginsberg. London: Methuen and Co. 1953. Pp. 82. 6*s*.

Prof. Ginsberg has surveyed the history of the doctrine of human progress from its beginnings in the eighteenth century up to the present day, and has given, as might have been expected, a clear and concise account of the many forms that the belief in progress has taken down the years. At one time progress was believed to be a moral advance towards human liberty and equality, and it was believed that the altruistic tendencies in human behaviour were gradually and by nature becoming more and more prominent. These beliefs could hardly stand up to even the least searching scrutiny by anybody who looked around him to-day, and modern writers like Bury, Inge and Baillie have raised doubts that will probably never be resolved about such theories of progress.

Prof. Ginsberg shows that neither development nor evolution, real as these processes are, need be progress, or even a condition of it, and he is doubtful of the religious solution of the problem given, e.g. by Baillie, on the ground, mainly, that all mature religions have much the same ethical content and that no one religion can justly claim to have the right to supersede others.

The conclusion to which Prof. Ginsberg comes is that there are trends in history which can be adjudged progressive and that by examining the conditions of success and failure, we can, if we choose, move further in the direction indicated by them. There has been a development in knowledge of human relations and in the effort to organize them, and to make actual the possibilities revealed by such knowledge and effort is to follow in the line of progress. It is, of course, not certain that advancing knowledge will be turned to greater social good, as used to be assumed, but, if rationality is taken seriously, there is a correlation between intellectual and moral development. There is at least 'no ground for suggesting any permanent disharmony between social justice and social efficiency. The choice therefore lies with ourselves.'

Prof. Ginsberg's essay is based on a lecture, and even in this extended form, it is highly compressed and very closely reasoned. It is not light reading, and a further expansion at many points would have made it clearer and more telling, but attentive study of what he has given us will be found rewarding.

GEORGE BROWN

Contributions of Psychology to Social Problems. By Cyril Burt. London: Oxford University Press. 1953. Pp. 76. 5*s*.

Previous Hobhouse Memorial Trust Lecturers have set a high and exacting standard. As those who were fortunate enough to hear it delivered already know, Sir Cyril Burt easily and gracefully maintains this standard in his memorial to Hobhouse—the twenty-second in the series—now published complete with

a wealth of footnotes and references. Taking as his starting-point Hobhouse's initial postulate that 'society is fundamentally a psychological structure', Sir Cyril proceeds to review the contributions of later psychologists, theoretical and applied, to the new and far-reaching questions which Hobhouse, perhaps the first Gestaltist, raised as a result of his own studies and inquiries.

In traversing this immense border country where the writs of psychology and sociology almost wholly overlap, Sir Cyril has no difficulty in showing that the contribution of the psychologist is an impressive one, most of all perhaps to the theory and practice of education. But this does not lead Sir Cyril into the arms of complacency. He has much to say that is both wise and timely about some of the dangers in current psychological researches, about neglected areas of inquiry, and about the need for the study of behavioural manifestations as distinct from verbal. He is still as intensely curious about all things human as he was when, in studying delinquency more than thirty years ago, he managed to get himself accepted as a member of a West End gang of criminals. R.M.T.

Psychology. By Ross Stagner and T. F. Karwoski. New York, Toronto and London: McGraw-Hill Book Co. Inc. 1952. Pp. xiii+582. 42*s.* 6*d.*

Stagner and Karwoski have attempted to write a text-book which does not avoid the more controversial issues at the expense of adhering too closely to the factual level. They feel that, since the student will in any case attempt to form an integrated view of behaviour, it is better not to leave this process to chance. They have accordingly organized their material around a guiding principle, the concept of homeostasis.

While this procedure will not commend itself to all teachers of psychology, it seems to the present reviewer that the authors must be held to have succeeded in providing a notably coherent and dynamic view of behaviour, which at the same time shows no marked theoretical bias. They have neither glossed over controversial matters, nor discussed them at a level beyond the grasp of the beginning student. The illustrative material has been chosen with discrimination from a wide range of sources. The book as a whole gives a very fair picture of contemporary psychological research and theory, and merits serious consideration as an introductory text. N.E.L.

The Development of Human Behaviour. By Richard Dewey and W. J. Humber. New York: Macmillan. 1951. Pp. xiv+762. No price given.

The title of this book is misleading: though some attention is given to the psychology of age groups, the book as a whole is intended to be an introduction to social psychology. It has its origins in a course in that subject given jointly by a sociologist and a psychologist, who rightly feel that social psychology would have much to gain from a better integration of the concepts of their respective disciplines than has yet been achieved.

More dubious is their proposal that 'social psychology is not a part of psychology or sociology, but a field of study in its own right'. The remainder of psychology they call 'physiological psychology', which leaves it unclear what is to become of the molar general psychology which studies behaviour without reference to its neural basis, or to any particular aspect of the environment. Whether or not Dewey and Humber regard this as 'physiological psychology', the first section of their book is devoted to a discussion of learning, motivation, and other topics which have most usually been studied in this way.

The account they give suffers considerably from over-compression and from the arbitrary selection of congenial theories. While the authors disavow any claim to finality for their account, their rather cursory and dogmatic presentation seems a little out of place in a book which claims to give the student a reliable view of human nature.

The central portion of the book is mainly concerned with the adjustment (or lack of it) between the roles which the individual can play and the status assigned him by his culture. The second section deals with the psychology of various age groups, more emphasis than usual being placed upon what behaviour the culture defines as appropriate to them. A third section discusses the psychology of minority status, and goes beyond the usual treatment of racial minorities to include, for example, the physically handicapped, the mentally ill, and the criminal. A final section describes behaviour in relation to social institutions, educational, religious, and economic.

These latter topics, in particular, are very large for the space which can be allotted to them, and there is some tendency to return to the rather didactic and sweeping manner of the introductory section. The treatment is generally sound in a descriptive way, but one could wish at times for more empirical material.

This book is none the less a creditable attempt to integrate data from diverse disciplines. One questions,

though, its suitability as an introductory text. Its first section gives rise to the feeling that the teaching of social psychology is better accomplished by the more traditional method of basing it upon an adequate course of 'physiological psychology'. The later discussions, too, might be of more value to students who already have some acquaintance with sociology and social psychology at the empirical level. N.E.L.

Psychological Theory. Arranged and edited by Melvin H. Marx. New York: The Macmillan Co. 1951. Pp. xi+585. 37*s.* 6*d.*

This book brings together a number of previously published papers by contemporary psychologists. Those in the first part are mainly on problems of theory and method, those in the second are by well-known psychologists on aspects of psychology to which they have made a special contribution. The editing has been kept to a minimum, but each chapter, which contains from two to eight papers on a common theme, is preceded by a brief paragraph summarizing the main purport of the papers. The book is intended for the use of advanced undergraduate or postgraduate students and should be particularly useful to those who have difficulty in obtaining access to journals and original sources. The editor has shown considerable judgement in the choice and arrangement of the papers, and he has seen fit to include among the papers on theory construction contributions ranging from those concerned with logical positivism to discussions of field theory and psycho-analytic postulates.

The first three hundred and fifty pages of the book are devoted to theory construction and reflect the intense contemporary interest in theory, but in spite of the clarity and the quality of the exposition the general impression left on the reader is of an over-preoccupation with minute questions of meaning. The finer differentiations between the subclasses of the intervening variable have a tendency to become niggling and fruitless. The discussions give no suggestion that psychology is even remotely concerned with human behaviour; at the most an occasional white rat makes a brief and apologetic appearance. The present preoccupation with the theory and methods of physics seems in danger of obscuring the need for investigations with a psychological content and meaning. It is as well to remember that logical positivism did not precede developments in the physical sciences; there is no historical reason for supposing that concern with semantics and the logic of experiment will be responsible for major developments in psychology.

The second and shorter half of the book is more recognizable as psychology and more concerned with human behaviour. For the undergraduate student it would provide a useful summary of some of the principal viewpoints in psychology and a brief introduction to important contemporary figures. He may also profit from the references at the end of some of the papers and the suggestions for further reading given at the back of the book. The average undergraduate would probably be bewildered by the first half of the book and the average graduate bored by the second half, but the book as a whole is of real interest largely because it gives in one volume so comprehensive a picture of psychology to-day. If the picture is a little dreary we can hardly blame the editor. M.M.L.

Essays in Psychology. Dedicated to David Katz. Uppsala: Almqvist and Wiksells. 1951. Pp. x+284. No price given.

This volume of eighteen essays was dedicated to the late Prof. Katz on his retirement, unfortunately to be so short. Fittingly, it is mainly the work of European psychologists, Carmichael and Thurstone being the only American representatives. Of the eighteen contributions ten are written in English, five in German, and three in French.

The distinction of the contributors, which include Bartlett, Burt and Thouless from this country, and such eminent names as Gemelli, Michotte, Piéron and Révész from the continent, is in itself enough to recommend the book. The material inevitably lacks unity, and is very various in subject-matter and treatment. The more important contributions are those which are not merely summaries of previously published work, but reports of unpublished investigations. Particular mention must be made of Michotte's essay, 'La perception de la fonction "outil"', which is an important appendix to his book *La perception de la causalité,* and of Johansson's account, which will delight the experimental psychologist, of his experiments on 'The effect of uniform and continuous chromatic changes'. Gemelli's essay on 'The effect of illusory perception of movement on sound localization' is also a significant experimental contribution. In a rather different field Thurstone's article on 'Experimental Tests of Temperament' is a valuable preview of recent Chicago work on the objective testing of temperament. Also in the field of temperament Ekman argues in a long essay that Sheldon's three-dimensional typology can be expressed

in terms of two dimensions. Burt provides some interesting historical material as background to the development of factor analysis. Révész discusses the place of gifted women in society.

Perhaps if any one thread is apparent in these diverse contributions it is an interest in the time dimension. This intrudes in most varied contexts; as anticipation in skilled performance with Bartlett; in the patterns of the dance in Buytendijk's 'Psychologie des Tanzes'; in Meyerson's philosophical reflections on the essential characteristics of human as distinct from animal behaviour. It is the central theme of Piéron's essay, 'Des aspects réels du temps en psychophysiologie', and is involved in Utitz's interesting article 'Zur Psychologie des provisorischen Daseins', in which he discusses psychological adjustment to modes of life known or expected to be temporary or provisional, e.g. holiday periods, or detention in a concentration camp.

There is considerable danger of articles published in volumes of this kind getting buried and forgotten. It would seem particularly important that they should be individually indexed in works of reference such as *Psychological Abstracts* and *L'Année Psychologique.* L.S.H.

Perception and Learning. By Olof Magne. Uppsala: Appelbergs Boktryckeri AB, 1952. Pp. 228. No price given.

The subtitle of this monograph is 'Experimental Studies in the Similarities and Dissimilarities between Perception and Learning Tests'. The studies are arranged as a doctoral dissertation in the University of Gothenburg. Detailed descriptions of the tests are given in appendices.

The 'leit motif' of the several investigations reported is associated with some of the earlier claims of Köhler that the same laws are valid for both perception and memory. For the present inquiry, *perception* is defined as 'a situation that comprises a reproduction or a statement in connexion with or immediately after a tachistoscopic exposition'. *Immediate retention* is defined as 'reproduction which is immediately preceded by a protracted exposition of stimuli, and which is constituted in such a way, that the organization of what is experienced need not necessarily take place in the same manner as in a perceptive situation, and reproduction is carried out in such a way, that as far as possible, sensory after-effects are prevented'. *Delayed retention* is described as 'the retention which can be found when a considerable time has elapsed after learning has taken place'.

The general finding from the several investigations is that, 'mental processes in learning situations are either rarely or never identical with the perceptive processes'. Evidence for this conclusion is derived from five main sources. The differential influence of some well-known 'effects' was studied. The 'von Restorff Effect' and the 'Katona Effect' were found to be more typical of learning than perceptual situations. The 'Wulf Effect' (the original claim was that of increasing symmetry and Prägnanz with successive reproductions) gave an inconclusive differentiation. There are, incidentally, some interesting references to Elmgren's failure to confirm the 'Zeigarnik Effect'. The study of a case of amnesia in a 15-year old boy following encephalomeningitis revealed a marked impairment in learning and recognition; but in 'perceptive' tests, his performance did not deviate significantly from that of children of his own age. Again, factorial analyses indicated that tests involving reproductions from long exposures had different factor loadings from those allowing short exposures.

The reader may feel that the distinction between perception and learning and the role of perception in learning has been treated somewhat arbitrarily; but the author has used some novel methods of investigation. F.V.S.

The Ames Demonstration in Perception. By William H. Nelson. Princeton, New Jersey: Princeton University Press; London: Geoffrey Cumberlege, Oxford University Press. 1952. Pp. xvi+88. 25*s.*

Some of the Ames demonstrations will be familiar to many in this country, in substance if not in name. They include the distorted room demonstrations and size-brightness demonstrations in which balloons may be varied in size by inflation and illuminated at different levels to show the influence of brightness and size in the judgement of distance.

The demonstrations have been described in other publications but up to the present there has been no single source in which the complete set of demonstrations—twenty in all—are even listed. The present publication fills this gap. Each demonstration is treated separately as a piece of laboratory equipment in such a way that anyone with workshop facilities could construct the apparatus without a great deal of

difficulty, although in some cases at a fair cost. The use of the apparatus is described and an outline is given of typical observations obtained.

The author emphasizes that no attempt is made to explain the phenomena that are experienced during the demonstrations. Various interpretations have been put forward in other publications and these are listed in a bibliography containing forty-nine references. The text is very fully illustrated with seventy drawings and photographs.

This publication will be invaluable to psychology departments and to everyone interested in visual perception. It is, in itself, an example of orderly presentation and the use of visual demonstration.

C.E.M.H.

The Acquisition of Word Meanings. By Heinz Werner and Edith Kaplan. Monographs of the Society for Research in Child Development, vol. xv, serial no. 51, no. 1, 1950. Evanston, Illinois. 1952. Pp. viii+120. No price given.

This is a striking and lamentable example of waste of energy. The authors set out 'to investigate experimentally the processes underlying the acquisition of word meaning through verbal contexts' in children. Their intention was to observe the changes in the manner in which children at successive ages respond to an unfamiliar word in a series of familiar verbal contexts. The authors rightly felt that the suggestion of Ogden and Richards that 'we early begin to use language in order to learn language' could provide a topic for an investigation that would make a contribution to our scanty knowledge of semantic development in children.

They devised a series of twelve tests, each consisting of six sentences, all including the same artificial word. For instance: for the word *corplum* ('stick, or piece of wood'); (1) A corplum may be used for support. (2) Corplums may be used to close off an open space. (3) A corplum may be long or short, thick or thin, strong or weak. (4) A wet corplum does not burn. (5) You can make a corplum smooth with sandpaper. (6) The painter used a corplum to mix his paints.

The tests were given individually to 25 children in each of the five years between 8 years 6 months and 13 years 5 months. From the results, presented with all due statistical precautions, the authors proceed to an elaborate and detailed discussion of such topics as the development of 'processes of signification, formation of verbal contexts, the ways meanings are given to an artificial word'; and find it necessary to use such terminology as 'protoforms of subsumptive generalisation'.

Now it is clear that everything depends on the tests. If the child is to arrive at the meaning of *corplum* from its six contexts, then the word must, at least, fit these contexts both semantically and grammatically. This is unfortunately not the case in all the tests. For instance: '*Soldeve* (wither, fade); '(1) The dinner was good but the fruit we ate was soldeve. (2) When we were driving in the evening we did not feel safe because things on the road seemed to soldeve. (3) The older you get the sooner you will begin to soldeve. (4) People like a blossoming plant better than one that is soldeve. (5) Putting the dress on the sunny lawn made the color of the cloth soldeve. (6) Because the windshield was frozen things looked soldeve.' Other examples with similar shortcomings could be cited.

As a result it becomes impossible to say how much is due to the immaturity of the children and how much to the invalidity of the tests. One can only hope that other workers will profit by the fundamental mistakes of this investigation.

M. M. LEWIS

The Psychology of Thinking. By W. Edgar Vinacke. London and Toronto: McGraw-Hill Book Co. Inc. 1952. Pp. xiii+392. 47*s.*

The investigation of the processes as distinguished from the topic of thinking is an intrinsically difficult task. This fact is probably responsible for the considerable literature on the experimental study of the thought processes and the relatively small knowledge about them. Prof. Vinacke, as he affirms in his preface, has attempted 'to survey in a systematic fashion the variety of human thought processes, together with what the psychologist has found out about them'.

There are chapters on 'Consciousness and the Field of Attention', 'Ideas, Imagery, and Imageless Thought', 'The Mechanism of Thought', 'Concept Formation', 'Creative Thinking', 'The Transfer-Retroaction Sequence', 'Problem Solving' and a welcome attempt in the last three chapters, 'Generalized Aspects of Personalized Thinking', 'Attitudes' and 'Public Aspects of Personalized Thinking', to relate the more formal study of the thought processes to later studies of personality, prejudices, and rumour.

The work is well indexed and documented and any student contemplating an experimental project in this or related fields will find in it useful references to much of the work done.

From the reader's point of view, the persistent attempt to survey gives place too infrequently to penetrating appraisal and analysis to adumbrate new developments; but a survey was the author's intention and a wide survey has been provided. F.V.S.

Straight and Crooked Thinking. By ROBERT H. THOULESS. London: Pan Books. 1953. Pp. 189. 2*s.* net.

Some people become psychologists because they think that psychology should have a contribution to make to the solving of human problems, individually, nationally and internationally. They are often disillusioned. Luckily Dr Thouless appears to have kept his illusions, and we should all be grateful. In this little book he explores one of the few applications of psychology where there is general agreement on principles, and where the application is of greatest importance in human affairs. The book is a revised and enlarged version of the 1930 edition. 'It exposes many dishonest tricks that are often used in argument'; and if therapy is at all possible through reading such a book, the patient might be expected to develop some immunity from the propaganda and 'tabloid thinking' of everyday life. The tricks which Dr Thouless classifies are sometimes psychological, more often logical. (Some of them are indispensable to advanced lifemanship, but the counterploys have thoughtlessly been provided.) Although Dr Thouless pokes fun at the detachment of the don, at times the book seems somewhat academic and humourless for its intended popular audience. (It is instructive to compare the book with the sparkling literary approach to a similar problem in Graves and Hodge's *Reader over your shoulder.*) However, it can be highly recommended as a discussion book for adult education classes. The reviewer has used it for this purpose with great satisfaction. BRIAN FOSS

A Textbook of Psychology (*Essential Facts and Views*). By S. JALOTA. Benares: Benares Hindu University Press. 1952. Pp. xvi+417. Rs. 12.

Prof. Jalota, encouraged, as he tells us, by the success in India of his Intermediate text-book, has written another, this time for Degree students. One wonders whether the author has not perhaps attempted too much. The topics covered range from the body-mind problem to parapsychology, and the theoretical position is so eclectic as to endanger the achievement of any coherent picture of the organism. The main influences appear to be those of Stout and McDougall and there is little reference to work published later than about 1930. N.E.L.

Range of Human Capacities. By DAVID WECHSLER. London: Baillière, Tindall and Cox. Second edition, 1952. Pp. x+190. 31*s.* 6*d.*

A second edition of this book is welcome because it should be read by all who are interested in psychology whether as professionals or as amateurs. Those who regard psychology as a biological science which should be planted firmly upon solid natural earth will read it to their comfort. Those who think otherwise will read it to their admonition. The task is not onerous because the book is very well written in a lucid and entertaining style free from technicalities.

The early chapters are a spirited sketch of problems involved in measuring psychological capacities, and stress is laid on the importance, neglected by some current psychometric methods, of making measurements in physical terms. Most of the remaining chapters are concerned with showing by means of data collected from many sources, that the ranges from highest to lowest capacity measured in these terms are small, varying from one broad class of human measurement to another, but within any one class surprisingly similar. The measures of range and their variation suggest an approach to the problem of the status of various capacities which is not developed but is of considerable interest. Two chapters, one on length of life and a very courageous chapter on some aspects of ageing, digress a little (as the author admits) from the main theme, but are well worth reading for their own sake.

Dr Wechsler has deliberately set out to be provocative—perhaps even a little waggish in places—and has certainly succeeded. No one who reads the book will fail to be stimulated by the views expressed. They may not always command the reader's assent, but they will certainly sharpen his critical eye and help to discipline his thinking.

The main regret is that the book is so expensive—a complaint made also by the reviewer of the first edition for this *Journal* seventeen years ago. A.T.W.

An Orally Presented Group Test of Intelligence for Juniors. By J. Cornwell. London: Methuen and Co. 1952. Pp. 40. 3*s.* 6*d.*

This is a one-hundred item test devised for children between the ages of eight and eleven. It consists of five subtests (same-opposite; true-false; problems; 'odd one out'; instructions); which are presented orally after blackboard examples. The children write the answers on ordinary exercise-book paper. In subtests 3 and 4 (multiple choice items) only the initial letter of the answer is to be written. Each item is read twice, and 10 seconds allowed before the next item is read. Each subtest is intended to be given 'on a separate occasion in a normal time-table period'.

The advantages of an orally presented test are obviously great, particularly to the teacher. The reviewer remembers how useful Spearman's 'Measure' was in the early thirties. Although Mr Cornwell compared the results from oral and written presentation of the test, unfortunately he gives no details of the result. Nor are there any data on the effects of change of expression, intonation or emphasis on the part of the tester. The test has been standardized on a group of 4150 children (8–12 years). Unfortunately, only the 10–11 age group ($N = 936$) was used to derive the norms, the table then being extrapolated up to 12 years and down to 7·6 years. It would have been preferable to have obtained the least-squares line from the whole sample, using Lawley's method. The norms give standard scores with a mean of 100 and S.D. of 16·5 (not 15, as is more customary with British group tests). The reliability (split-half) is 0·953, giving a standard error of quotient of 3·7. Correlations of 0·91 and 0·85 were obtained with Simplex and Terman-Merrill respectively.

This inexpensive test ought to be most useful to the class-teacher for internal grading. Difficulties may arise, however, if teachers attempt to compare the I.Q.'s from this test with results from, say, Moray House tests having a different basis of standardization. May we hope that Mr Cornwell will pursue 'oral *versus* written' presentation a little further, and perhaps issue revised norms when this becomes possible?

S.W.

Personality Measurement. By Leonard W. Ferguson. New York, Toronto, London: McGraw-Hill Book Co. Inc. 1952. Pp. xv+457. 51*s.*

Dr Ferguson presents a concise and lucid critical evaluation of a wide range of measurement techniques. His touchstones are reliability, validity, objectivity of scoring, feasibility of the method and usefulness of results. A particularly valuable feature of his approach is his use of the historical method in showing how early and unsatisfactory techniques were refined and improved by the rigorous application of statistical procedures, but he does not fail to note cases of tests which have remained unsatisfactory because the statistical procedures used in their refinement were based on bogus logic.

Amongst the measures discussed are the Strong Vocational Interest Test, the Kuder Preference Record, Thurstone's Attitude Scales, the Likert and Guttman techniques, the Bernreuter Personality Inventory, the M.M.P.I., Burgess and Cotterell's and Terman's studies of Marital Happiness, Kurtz and Kornhauser's Insurance Selling Aptitude Index, Industrial Merit Rating methods, the Forced Choice Rating technique, the Rorschach test, Murray's T.A.T., Lewin's observational approach, the O.S.S. Assessment Staff's work, and Eysenck's early work reported in *Dimensions of Personality.*

Dr Ferguson does not pull his punches and is not afraid to attack what he regards as big myths. The Rorschach cult is hammered for its 'crude and shallow analogies'—'If the Rorschach test has merit, very little of this merit has been demonstrated... because most Rorschachers back away from a rigid validation procedure'. The O.S.S. Assessment Staff's work is held up as an example of shocking bungling. Murray's T.A.T. is rejected as both dubious in its results and excessively inconvenient in its application. Amongst tests which receive approval are the Strong Interest Test, the M.M.P.I., the Insurance Selling Aptitude Index, Thurstone's Attitude scales and Eysenck's experimental approaches.

These facts seem to the reviewer to be sufficient evidence of the book's acceptably high level of reliability, validity and objectivity.

J. D. HANDYSIDE

The Influence of Social Class on Tests of Personality. Drew University Studies No. 5. By Frank Auld Jr. Madison, New Jersey: Drew University. 1952. Pp. 18. Free on application to Registrar's Office, Drew University.

The author shows that lower-class and middle-class children differ considerably in their responses to an inventory intended to assess degrees of maladjustment, and he gives reasons for thinking that the different scores reflect class-induced habits and attitudes rather than differences in personal adjustment.

He questions, therefore, the interpretation of some of the class differences found in Rorschach scores, and he rightly emphasizes the importance of keeping social status constant in assessing the influence on test responses of some other variable differentiating two groups. Although slight in itself this short paper raises a problem of major importance.

Controlled Projection for Children. By J. C. Raven. London: H. K. Lewis, 2nd edition, 1951. Pp. 176, with 4 plates (2 coloured). 25*s*.

The Individual's Relation to Reality in Our Culture. By Elizabeth F. Hellersberg. Springfield, Illinois, U.S.A.: Charles C. Thomas; Oxford: Blackwell. 1950. Pp. x + 128, with 74 illustrations. 25*s*.

A Test of Family Attitudes. By Lydia Jackson. London: Methuen. 1952. Pp. 37, with 8 test pictures. 10*s*. 6*d*.

It is interesting to study these three projection tests, and to compare the techniques and materials used and the results expected from them. Raven's controlled projection test, as we know from the first edition in 1944, depends on the use of free drawing in conjunction with a series of prearranged questions suitable for children. The test forms a useful method of research into the social attitudes, habits and personal relationships of children, and for examination of children for clinical purposes. It has been widely used and a number of interesting papers have been published about its application. The present edition is amplified in useful ways and contains some illustrations of children's drawings, beautifully reproduced, two of them in colour, tabulated responses of 150 children of ages $6\frac{1}{2}$, $9\frac{1}{2}$ and $12\frac{1}{2}$ years, and a great amount of other valuable material obtained by the application of the test.

The projection test described by Hellersberg depends on the completion of pictures in 12 'blanks' in which certain standard preliminary lines are offered, and on making a drawing in one completely blank space. This technique is amplified by careful questioning of each subject about the object drawn and his attitude to it, his interest in it, where he saw it and so on, and on careful recording of all the observations the subject makes and of his answers to the questions. Here, as compared with Raven's Test, the drawing is more controlled, but the questioning more free. The aim of this test and of the technique of interpretation is to gain some estimate of the kind, nature and degree of the subject's realtiy-contacts in daily life, and it should be useful to clinical psychologists.

The family attitudes test published by Jackson consists of a series of seven pictures suitable for children, which bear upon situations of possible stress to a child in the family circle. The second picture is given in two forms, one for boys and the other for girls. The other pictures are considered suitable for either sex. The child is invited to make up a story about each picture, and the story is carefully recorded. Clear instructions are given in the book, together with a set of standard questions to be asked about each picture. Experience with this test before its publication has convinced the reviewer that it is a valuable addition to the tester's and therapist's battery of tests. Regrets may be felt that it is limited to family attitudes.

Although it might be said that the essence of projection tests lies in freedom, all three of these tests use control or restriction to some extent or in some way.

The student of tests may spend some time profitably in considering the relative claims of partly controlled or restricted projection tests like these and the traditional free projection tests such as those of Rorschach and Murray. What is lost in depth of penetration into the personality and its unconscious conflicts may be gained in speed, ease and rapidity of interpretation and in clarity of result, but the illumination of the inner qualities and aspects of the subject's personality suffers at the same time to some extent.

R. W. PICKFORD

Primary Love and Psycho-Analytic Technique. By Michael Balint. London: Hogarth Press. 1952. Pp. 265. 30*s*.

The publication of Dr Balint's collected papers (together with some admirable contributions by his late wife, Alice Balint), will be welcomed by all who are interested in psycho-analysis. The first part of the book, entitled Instincts and Object-Relations, is devoted to an examination of the underlying basis of love and hate, while the second is concerned with Problems of Technique. Dr Balint's ideas are particularly important since he is at pains to re-affirm the biological basis of his work while at the same time possessing a deep awareness of the psychological implications of human relationships—that is to say, he approaches

the problems of the psyche from the directions both of physiology and sociology. (In the preface, he acknowledges his debt to his wife and draws attention to her interest in anthropology which contrasts with his own in the more exact sciences. The valuable fruits of this partnership are visible throughout the work.) This is of particular significance because many modern psycho-analysts, impressed by the relevance of the social sciences, have forgotten the existence of the biological basis so emphatically stressed by Freud. But Dr Balint recognizes mind as being, at least in part, a product of the tension of the physiological and the social. Without explicitly being intended as such, his book is a significant contribution towards the reconciliation (I do not say the integration, for that may be an unrealizable aim) of the different sciences of behaviour. No set of general theories emerge, but the whole orientation of the writer is convincingly and brilliantly established in a long succession of profound and erudite discussions of specific issues.

But we must gently take Dr Balint to task for following precedent in publishing collected papers delivered over a long period—in his case, of twenty-two years. Dr Balint says that his ideas reached a 'certain completeness' in the last two papers. In addition, it is perfectly true that the other papers, excellent in themselves, give a picture of the development of his ideas which is extremely interesting. But how valuable it would be if eminent analysts of Dr Balint's maturity could spare the time to give us a coherent and connected account of the fruits of their experience. A.C.

Success in Psychotherapy. Edited by W. Wolff and J. A. Precker. New York: Grune and Stratton. 1952. $4.75.

This book contains seven papers, of which the first is 'Problems in the definition and measurement of success in Psychotherapy' by Mosak. The concept of success is criticized as a value judgement and intrinsically subjective. The users of the 'sign' approach are then castigated for their claims to 'global evaluation of personality'. A method should not, however, be dismissed because it is sometimes abused in this way. Having thus cut the ground from under his own feet, the author can only end by trying to improve the quality of the bootlaces by which he has perforce to endeavour to lift himself.

In 'Objective measurement in evaluating psychotherapy', Thetford measures the level of autonomic activity in relation to frustration in a group of subjects before and after psychotherapy. Those who expressed an ambivalent attitude to the outcome of treatment showed a more marked change in the physiological measurement than did those expressing a positive or a negative attitude. This otherwise useful experiment is somewhat marred by the curious belief that 'the physiologic disequilibrium, although it may not manifest itself in observed behavior, is a more basic indicator of the threshold for frustration tolerance'.

Each of the next three papers contains at least four of the following weaknesses: (*a*) faulty use of statistics, (*b*) inadequate control groups, (*c*) ratings of demonstrated low reliability, (*e*) clinical naïvety, (*e*) specious argument, (*f*) reiteration, (*g*) tautologizing, (*h*) ponderous platitudes, (*i*) unnecessary verbiage. A few examples, using the above letters for identification, must suffice.

(*a*) Hogan writes: 'As represented by lack of significant difference between judgments, agreement was perfect for judgments of the amount of defensiveness in each interview and for the case as a whole.'

(*c*) Haimowitz and Haimowitz use ratings with a mean correlation coefficient of 0·53. A judge was trained when he achieved 60% agreement with Hogan.

(*e*) Haimowitz and Haimowitz attach importance to a difference of one neurotic sign on the Rorschach test even though this seems to be due mainly to the disappearance of colour and shading shock on re-test, when subjects would presumably be expecting shaded and coloured cards.

(*f*) In an article of less than eighteen pages Hamlin *et al.* state their results four times, apart from tables.

(*g*) Hogan states that 'defensive behavior may be identified by...identifying a defensive form of behavior or a defensive manner of behaving'.

Raskin's work on the change after psychotherapy in the patient's 'locus-of-evaluation' of himself and of his dependence on the judgements of others might be regarded as rather more successful, but as little more than a pilot survey requiring some overhaul.

H. H. Anderson, in a much more readable paper, discusses the hypothesis of the 'Growth Circle', which is that integrative behaviour in one person tends to induce integrative behaviour in others. The patient's ability to break out of the vicious circle of domination that incites domination in others and to develop a more integrated pattern of behaviour he sees as a useful criterion of successful therapy.

The book can be recommended to those psychiatrists who distrust the 'scientific' approach of the psychologist. G.A.F.

The Case Against Psycho-Analysis. By Andrew Salter. London: Medical Publications Ltd. 1953. Pp. ix+179. 10*s.* 6*d.*

This is a wearisome book. It is monotonously insistent on its central thesis that there is nothing of value in any aspect of Freudian theory or therapy. The author has, moreover, carefully included all the more commonplace and stereotyped misunderstandings of psycho-analysis. A heavy loading of emotionally toned words further reduces what slight value, as a contribution to popular education, the book might yet have had. 'Erroneous obscenities' (p. 99), 'old-fashioned sexual monomania' (p. 133), 'profane fictions' (p. 62), 'preposterous' (a favourite word), and 'complete idiocy' (p. 60)—a serious branch of modern psychological science cannot be dismissed in such terms as these. Unfortunately, this is the kind of book which will strongly reinforce both the sentiments and the misunderstandings of the already adversely prejudiced. It is unlikely to be taken very seriously by those whose attitude towards modern psychology is scientific. The absence of an index in a book of this kind is to be expected.

PETER McKELLAR.

Prescription for Rebellion. By Robert Lindner. London: Gollancz. 1953. Pp. viii+305. 16*s.* net.

This high-pitched protest against the practices of psychiatrists and psycho-therapists when they aim simply at the 'adjustment' of the patient to the world around him is illustrated with lurid examples from American experience. Dr Lindner's outcry can be seen as part of the wider dissatisfaction now being voiced by some psychologists against the loss of individuality in mass civilization. He asks for recognition 'That there exists an inherent right of rebellion based upon an innate biological drive' (p. 278), and his positive programme depends on a sweeping, though not very fully defined change in educational practice: 'I am asking that the first principle of education become one which recognizes the function of teaching to be that of assisting the process of self-understanding—awareness, that is, based upon the right of protest rather than the imperative of adjustment' (pp. 279–80). Dr Lindner gives too little attention to the content and purpose of the protest and seems almost to imply that individuality can develop regardless of the wealth or poverty of the cultural opportunities open to the developing mind. The shallowness of the cultural soil in mass civilization is a problem he neglects. But his chief need was to raise a clamour against the inhuman folly of taking social adjustment as the criterion of mental health, and this he has done.

The Psychology of Religion. By L. W. Grensted. London: Geoffrey Cumberlege, Oxford University Press. 1952. Pp. vi+181. 6*s.*

This new Home University volume surveys the contribution of modern psychology to the study of religion. The exposition is concise and clear and the author is to be congratulated upon his skilful inclusion of all the main topics in a volume of only 181 pages. Particularly happy is the balance he has struck between historical information and the discussion of living issues, as this will make the book of interest and value both to beginners and advanced students.

The seven chapters of the book deal with: the subject-matter of the psychology of religion; a descriptive account of what is meant by the term religious; a review of the psychological approach to religion; individual religion; the genetic approach in various schools of psychology; corporate religion in the light of social psychology and the discussion of reality from the standpoint of psychology. There is a short note on psychical research. Of special use to the beginner is the bibliographical note at the end.

Teachers and students alike have hitherto been hampered by the lack of an adequate guide to the study of the psychology of religion. To produce one which is both intellectually stimulating and simply written is an achievement. The book is much to be welcomed.

S. LAWSON JONES

The Psychology of Love in Forgiveness. By Ernest White. Spiritual Healing Booklets No. 4. London: Epworth Press. 1953. Pp. 23. 1*s.* net.

This booklet, written from the Methodist standpoint, draws attention to the relative neglect in psychology of the concept of forgiveness and the related topics of sulking and repentance. It is a sensible tract, intended only for the general reader, but it raises topics of which psychology might well attempt a more adequate account.

La Psychologie Différentielle. By H. Piéron. Livre 1e. Traité de Psychologie Appliquée. Paris: Presses Universitaires de France. 1949. Pp. viii+122. 400 fr.

This book is the first of some seven which together, when complete, will make up a comprehensive work on applied psychology, edited by Prof. Piéron. As such it calls for little comment beyond the statement that the Director of the Institut National d'Étude du Travail et d'Orientation Professionelle shows all the lucidity and erudition of his *alter ego*—the late Professor of the Physiology of Sensation at the Collège de France. The ultimate value of the whole work, to which other authors will contribute, will have to be assessed when more of it has appeared, but Prof. Piéron has set a high standard for it to live up to.

The present part contains three chapters, respectively devoted to the 'Role of Heredity', 'Aptitudes', and 'Individuality and the Problem of Types'. Each brings together a tremendous amount of information in a small space and within a traditional framework. References to recent work are given in plenty, but the general outlook is rather that of a continuing past than of an unfolding future. For the student it will form a useful work of reference.

Psychology: General, Industrial, Social. By J. M. Fraser. London: Pitman. 1951. Pp. x+310. 25*s.* net.

In writing this book the author was clear as to the readers for whom it was intended and the need it was to fulfil. He had in mind a recommendation of the Committee on Education for Management, which was that Psychology be included in the syllabus for a Common Intermediate Examination as a 'background' subject, and which advocated a course of study 'designed to provide an introduction to General Psychology and to illustrate its application in dealing with the human problems of industry and commerce and the management of social groups'. The book is so designed. In it the fields of psychology indicated in the title are surveyed not so much from the manager's viewpoint as from a viewpoint which the manager students can share and the main matters relevant to 'the art and science of handling men' are clearly explained and discussed. The book serves its purposes well and seems to the reviewer likely to be understood and appreciated by the students for whom it is intended. J. P. LEES

Selected Sources of Free and Inexpensive Information Concerning Vocational Rehabilitation —A Bibliography. By Lynn L. Ralya and Lillian L. Ralya. Dr L. L. Ralya, 907, Fourteenth Street, Santa Monica, California. 1953. Pp. 9. 25 cents.

This short list of pamphlets, mostly American and mainly for the information and education of the general public, includes sections on vocational rehabilitation in general, the ageing, the alcoholic, amputees, the blind, the cardiac and the circulatory handicapped, the cerebral palsied, the epileptic, the hearing handicapped, the infantile paralysis handicapped, the mentally retarded and the emotionally disturbed, the tuberculous. It testifies to the serious efforts being made in America to provide popular guidance on these problems.

OTHER PUBLICATIONS RECEIVED

ROGER W. RUSSELL. *The Comparative Study of Behaviour.* London: H. K. Lewis and Co. Ltd. 1952. Pp. 19. 4*s.* net.

FR. AGOSTINO GEMELLI and SAC. GEORGIO ZUNINI. *Introduzione Alla Psicologia.* 2nd edition. Milan: Società Editrice 'Vita e Pensiero'. 1952. Pp. xv+490. L. 1100 net.

JOHN P. FRANK. *My Son's Story.* London: Sidgwick and Jackson Ltd. 1952. Pp. xiv+174. 10*s.* 6*d.* net.

DONALD McI. JOHNSON. *Indian Hemp.* London: Christopher Johnson. 1952. Pp. 112. 8*s.* 6*d.* net.

DR MOHAMED HOSNY WELAYAH. *Some Case Records of an Egyptian Psychoanalyst.* Vols. I and II. Alexandria (Egypt): Aux Editions du Scarabée, 1951 (Vol. I); and Imprimeries Al. Bassir, 1952 (Vol. II). 10*s.* each.

GORDON WESTWOOD. *Society and the Homosexual.* London: Victor Gollancz Ltd. 1952. Pp. 191. 9*s.* 6*d.* net.

HENRY ENOCH KAGAN. *Changing the attitude of Christian toward Jew.* New York: Columbia University Press (Geoffrey Cumberlege). 1952. Pp. xvi+155. 18*s.* net.

JAMES T. FISHER and LOWELL S. HAWLEY. *Diary of a Psychiatrist.* London, N.Y., etc.: Medical Publications Ltd. 1952. Pp. 256. 16*s.* net.

MARION E. KALKMAN. *Introduction to Psychiatric Nursing.* N.Y., Toronto and London: McGraw-Hill Book Co. Inc. 1950. Pp. xvi+336. $3.75 net.

E. ORSON BROWER. *Should Fatherhood Begin at Forty?* Philadelphia: Elbert Orson Brower. 1952. Pp. xvi+240. $3.25.

LEE R. STEINER. *A Practical Guide for Troubled People.* New York: Greenberg, Publisher. 1952. Pp. 299. $3.50.

LUCY FREEMAN. *Fight Against Fears.* London: Victor Gollancz Ltd. 1952. Pp. 287. 12*s.* 6*d.* net.

ARNOLD HUTSCHNECKER. *The Will to Live.* London: Victor Gollancz Ltd. 1952. Pp. viii +278. 13*s.* 6*d.* net.

La Psychotechnique dans le Monde Moderne. Bibliothèque Scientifique Internationale (Ed. Dr Franziska Baumgarten). Paris: Presses Universitaires de France. 1952. Pp. xi+630. 2000 fr.

RAYMOND RUYER. *Philosophie de la Valeur.* Paris: Librairie Armand Colin. 1952. Pp. 215. (No price given.)

ANDREW SALTER. *Conditioned Reflex Therapy.* London: George Allen and Unwin Ltd. 1952. Pp. ix+358. 20*s.* net.

ABRAHAM MYERSON. *Speaking of Man.* London: Secker and Warburg. 1952. Pp. 255. 15*s.* net.

Prof. Dr Willy Hellpach. *Mensch und Volk der Großtadt.* Stuttgart: Ferdinand Enke Verlag. 1952. x+153. DM. 14.70.

Kauki K. Kaila. *Über den Zwangsneurotischen Symptomenkomplex.* Kopenhagen: Ejnar Munksgaard. 1949. Pp. 227. (No price given.)

Board of Visitors of Letchworth Village. *Research into the Causes of Feeblemindedness.* Utica (N.Y.): State Hospitals Press. 1952. Pp. 37.

Earle L. Reynolds. *The Distribution of Subcutaneous Fat in Childhood and Adolescence.* Evanston, Illinois: Child Development Publications. 1951. Pp. xviii+189. Monographs of the Society for Research in Child Development, Inc. Vol. xv, Serial no. 50, no. 2, 1950.

Kali Prasad. *Fatigue and Efficiency in the Textile Industry.* (No. 2, Laboratory of Experimental Psychology, Lucknow University.) Lucknow: Laboratory of Experimental Psychology. 1952. Pp. 46. 2*s.* 3*d.* net.

Current Sociology. No. 1, An international bibliography of sociology. Paris: U.N.E.S.C.O. 1952. Pp. 81. Each number, $1.00; 6*s.*; annual subscription, 18*s.*

Peter R. Hofstätter. *Die Psychologie und das Leben.* Wien-Stuttgart: Humboldt-Verlag. 1951. Pp. 288. (No price stated.)

Theses in the Social Sciences. Paris: U.N.E.S.C.O. 1952. H.M.S.O. Pp. 236. 7*s.* 6*d.* net.

What is Race? Department of Mass Communication of U.N.E.S.C.O. (Illustrations by J. E. Kleiman and text by Diana Tead.) Paris: U.N.E.S.C.O. 1952. Pp. 87. 5*s.* net.

International Repertory of Social Science Documentation Centres. Co-ordinating Committee on Documentation in the Social Sciences for the Department of Social Sciences, U.N.E.S.C.O., Paris. 1952. Pp. 42. 2*s.* 6*d.*